PATENT
Docket No.: 19603/3357 (CRF D-1595G)

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicants	:	Barany et al.	Examiner: To Be Assigned
Serial No.	:	09/986,527	
Cfrm. No.	:	To be Assigned	Art Unit: To Be Assigned
Filed	:	November 9, 2001	
For	:	DETECTION OF NUCLEIC ACID SEQUENCE DIFFERENCES USING THE LIGASE DETECTION REACTION WITH ADDRESSABLE ARRAYS	

DECLARATION OF FRANCIS BARANY UNDER 37 CFR § 1.608(b)

Commissioner of Patents
Washington, D.C. 20231

Dear Sir:

I, FRANCIS BARANY, pursuant to 37 CFR § 1.608(b), declare:

1. I received a B.A. in Chemistry in 1976 from the University of Illinois at Chicago Circle and a Ph.D. in Microbiology from The Rockefeller University in 1981.

2. I am a Professor of Microbiology at the Joan and Sanford I. Weill Medical College of Cornell University, New York, New York.

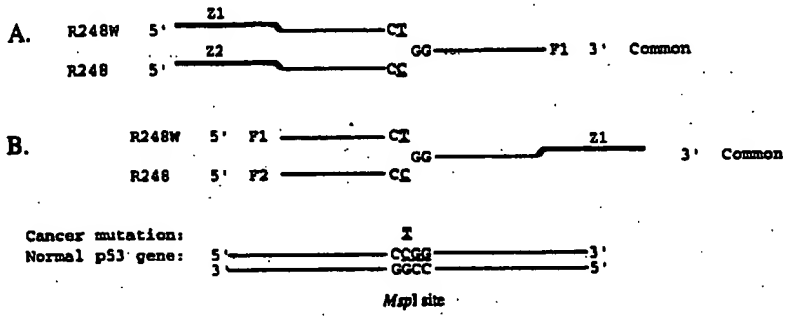
3. I am a named co-inventor of the above-identified patent application.

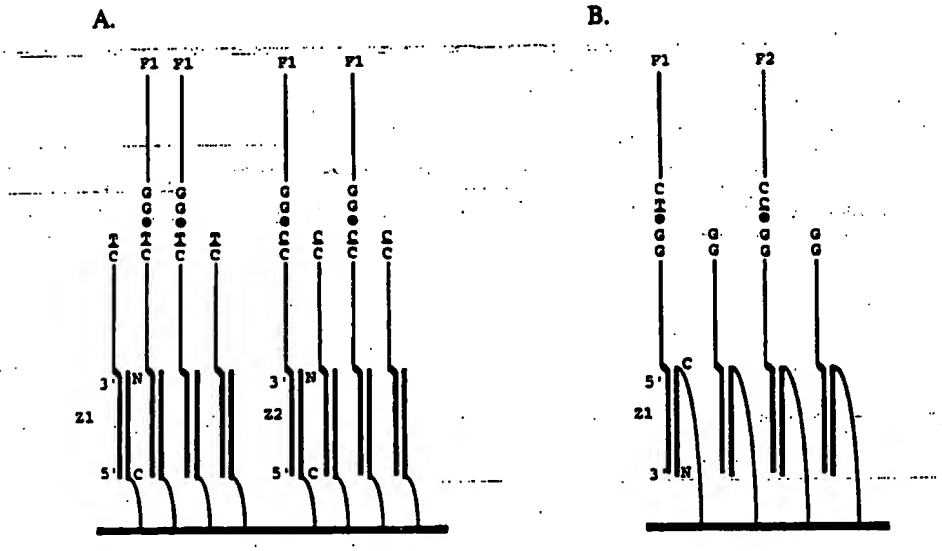
4. I am submitting this declaration to demonstrate that (1) on or before February 4, 1994, my co-inventors and I conceived the idea of using an array of a plurality oligonucleotide analogue probes coupled to a solid support where the probes bind to complementary oligonucleotide targets with similar hybridization stability across the array and (2) on or before May 31, 1994 to June 2, 1994, this conception was communicated by me and my co-inventors to an employee of Affymetrix, Inc., Santa Clara, California.

5. On or before February 4, 1994, a grant application, entitled "New Methods for Cancer Detection", ("Grant Application") was submitted to the National Cancer

Institute, U.S. Department of Health and Human Services ("NCI") by, amongst others, me, Donald Bergstrom, and my co-inventors George Barany and Robert P. Hammer. The Grant Application describes five projects, including Project 5, entitled "Design and Synthesis of DNA and PNA Arrays". A copy of relevant portions of the Grant Application, including the Project 5 description is attached hereto at Appendix 1. It is my understanding that the portions of the Grant Application which are attached at Appendix 1 are an accurate version of what was submitted to NCI on or before February 4, 1994.

6. Project 5 describes a method of detecting nucleic acid sequence differences in target nucleic acids by the use of a solid support with an array of peptide nucleotide analogues ("PNA") to capture and detect the products of a ligase detection reaction ("LDR") (Appendix 1, p. 19-22). In particular, each ligation product of the LDR process is provided with a "zip code" tail (e.g., Z1 and Z2) which is selectively captured by a complementary "zip code" located at a particular "address" on the solid support (Id.). As shown below, different target nucleic acids captured in this manner are detected and distinguished from one another by the presence of labels immobilized at different addresses on the array (having "zip code" capture probes with different nucleotide sequences) or by the presence of different labels immobilized on the array (Id. at 22).





R528437.1

The Grant Application discloses that the array "zip code" capture probes on the solid support can be designed to achieve optimal hybridization to the "zip code" tail on the ligation product (Id.). This is achieved by increasing thermal stability through narrowing of the difference in melting temperature (i.e. T_m) between the different duplexes formed by LDR products having a "zip code" tail and the complementary "zip codes" (i.e. PNA oligonucleotides) hybridized to one another on the solid support (Id.). These differences in melting temperature result from differences in G•C/A•T content (Id.). The Grant Application states that the T_m for correct PNA/DNA hybridization is at least 48°C higher than for any incorrect hybridization and neighboring 24-mer capture probes on the solid support are separated by 12-mer oligonucleotides which do not hybridize to anything (Id. at 31). The Grant Application also states that analogues with thymine replaced with 5-propynyl uridine can be used in either complementary zip codes (DNA or PNA) on the solid support or zip code (DNA) tails on the ligation product (Id. at 24). The subject matter of the claims of the present application are described in our Grant Application as set forth in the following paragraphs.

7. Claim 1 of the present application is directed to:

> A composition for analyzing interactions between oligonucleotide targets and oligonucleotide probes comprising:
> an array of a plurality of oligonucleotide analogue probes having different sequences, wherein said oligonucleotide analogue probes are coupled to a solid substrate at known locations and wherein said plurality of oligonucleotide analogue probes are selected to bind to complementary oligonucleotide targets with a similar hybridization stability across the array.

The subject matter of this claim is described in the Grant Application as follows. The Grant Application discloses an array of oligonucleotides on a solid support (Id. at 19, lines 36-44) and its use in a method for detecting nucleic acid sequence differences in target nucleic acids by hybridization of nucleic acid sequences to the oligonucleotides on the array (see, e.g., Id. at 11, lines 4-6; Id. at 16, lines 1-14). A plurality of DNA oligonucleotides or peptide nucleotide analogue (PNA) oligomers (complementary zip codes) are attached at different locations to a solid support, each having different nucleotide sequences (Id. at 21, lines 27-

30; Id. at 22, lines 9-10; Id. at 25, lines 5-8; Id. at 19, lines 36-44). Thymine can be replaced by 5-propynyl uridine when used within DNA or PNA address sequences (attached to a support) as well as in DNA zip code sequences (Id. at 24, lines 8-9) and PNA oligomers can be utilized in unique addressable arrays (Id. at 21, lines 27-30). Moreover, the Grant Application discloses a PNA oligonucleotide array with 1,296 addresses (Id. at 21, lines 53-54), a complementary zip code at a discrete address on a two-dimensional array (Id. at 35, lines 16-18), and LDR products having a "zip code" tail, which will be selectively captured by a complementary zip code (DNA oligonucleotides or PNA oligonucleotides) on a solid support (Id. at 11, lines 4-7; Id. at 19, lines 42-44). Zip code (oligonucleotide) capture can be optimized by narrowing the thermal stability (T_m) difference between duplexes formed by LDR products having a zip code tail and the complementary zip codes (DNA or PNA oligonucleotides) on the solid support hybridized to one another resulting from differences in G•C/A•T content (Id. at 22, lines 18-20). The Grant Application discloses raising the T_m for all zip code/address duplexes (Id. at 22, lines 18-24). The T_m for the correct PNA/DNA hybridization is at least 48 °C higher than any incorrect hybridization and neighboring 24-mers are separated by 12-mers, which do not hybridize with anything (Id. at 31, lines 23-25).

8. Claim 12 of the present application is directed to:

> A composition for analyzing the interaction between an oligonucleotide target and an oligonucleotide probe comprising:
> an array of a plurality of oligonucleotide probes having different sequences hybridized to complementary oligonucleotide analogue targets, wherein said oligonucleotide analogue targets bind to complementary oligonucleotide probes with a similar hybridization stability across the array.

The subject matter of this claim is supported by the disclosure of the Grant Application as follows. The Grant Application discloses an array of oligonucleotides on a solid support (Id. at 19, lines 36-44) and its use in a method for detecting nucleic acid sequence differences in target nucleic acids by hybridization of nucleic acid sequences to the oligonucleotides on the array (see, e.g., Id. at 11, lines 4-6; Id. at 16, lines 1-14). In addition, the Grant Application discloses an array of oligonucleotides on a solid support (Id. at 19, lines 36-44), a plurality of different sequences attached at different locations (addresses) on the solid support (Id. at 19, lines 36-44), and a universal PNA oligonucleotide array with 1,296 addresses (Id. at 21, lines 53-54). LDR products having a zip code tail hybridize to complementary zip codes on a solid support where the complementary components are DNA or PNA (Id. at 11, lines 4-6).

Thymine can be replaced with 5-propynyl uridine when used within DNA or PNA address sequences (complementary oligonucleotide probes) as well as in the DNA zip code sequences (Id. at 11, lines 8-10; Id. at 24, lines 8-9). Moreover, the Grant Application discloses that zip code (oligonucleotide) capture can be optimized by narrowing the thermal stability (T_m) difference between duplexes formed by LDR products having a zip code tail and the complementary zip codes (DNA or PNA oligonucleotides) on the solid support hybridized to one another resulting from differences in G•C/A•T content (Id. at 22, lines 18-20). In addition, the Grant Application discloses raising the T_m for all zip code/address duplexes (Id. at 22, lines 18-24) and discloses that the T_m for the correct PNA/DNA hybridization is at least 48 °C higher than any incorrect hybridization and neighboring 24-mers are separated by 12-mers, which do not hybridize with anything (Id. at 31, lines 23-25).

9. Claim 15 of the present application is directed to:

> A method of analyzing interactions between an oligonucleotide target and an oligonucleotide probe comprising the steps of:
>
> (a) synthesizing an oligonucleotide analogue array comprising a plurality of oligonucleotide analogue probes having different sequences, wherein said oligonucleotide analogue probes are coupled to a solid substrate at known locations, said solid substrate having a surface;
>
> (b) exposing said oligonucleotide analogue probe array to a plurality of oligonucleotide targets under hybridization conditions such that said plurality of oligonucleotide analogue probes bind to complementary oligonucleotide targets with a similar hybridization stability across the array; and
>
> (c) determining whether an oligonucleotide analogue probe of said oligonucleotide probe array binds to at least one of said target nucleic acids.

The subject matter of this claim is supported by the disclosure of the Grant Application as follows. The Grant Application discloses a method for detecting nucleic acid sequence differences in target nucleic acids using an array of oligonucleotides on a solid support to capture complementary zip code tails on LDR products (Id. at 11, lines 4-6). A plurality of DNA oligonucleotides or PNA oligomers (complementary zip codes) are attached at different locations to a solid support, each having different nucleotide sequences (Id. at 21, lines 27-30; Id. at 22, lines 9-10; Id. at 25, lines 5-8; Id. at 19, lines 36-44). Thymine can be replaced by 5-propynyl uridine when used within DNA or PNA address sequences (attached to a support) as well as in DNA zip code sequences (Id. at 24, lines 8-9). Moreover, the Grant

Application discloses PNA oligomers in unique addressable arrays (Id. at 21, lines 27-30), a PNA oligonucleotide array with 1,296 addresses (Id. at 21, lines 53-54), and a complementary zip code at a discrete address on a two-dimensional array (Id. at 35, lines 16-18). In addition, the Grant Application discloses a solid support having a surface (Id. at 21, lines 4-5, 31, and 48-49; Id. at 25, lines 9-12). Ligated reaction product is contacted with a solid support including complementary zip codes (DNA oligonucleotides or PNA oligomers) which are complementary to addressable array-specific portions (zip codes) of a ligated product sequence (Id. at 11, lines 4-7). The zip code tail of the ligated reaction product will be captured by a complementary zip code on the solid support (Id. at 11, lines 5-6). Moreover, the Grant Application discloses incorporation of the nucleotide analogue, 5-propynyl uridine, into DNA zip code and PNA address sequences (Id. at 11, lines 8-10) and discloses that zip code (oligonucleotide) capture can be optimized by narrowing the thermal stability (T_m) difference between duplexes formed by LDR products having a zip code tail and the complementary zip codes (DNA or PNA oligonucleotides) on the solid support hybridized to one another resulting from differences in G•C/A•T content (Id. at 22, lines 18-20). In addition, the Grant Application discloses raising the T_m for all zip code/address duplexes (Id. at 22, lines 18-24). Further, the Grant Application discloses that the T_m for the correct PNA/DNA hybridization is at least 48 °C higher than any incorrect hybridization and neighboring 24-mers are separated by 12-mers, which do not hybridize with anything (Id. at 31, lines 23-25). The presence of ligation products is identified by detecting labels of oligonucleotide probes hybridized to the solid support (Id. at 19, line 44, to 21, line 25; Figure 2).

10. Claim 25 of the present application is directed to:

> A method of detecting an oligonucleotide target comprising:
>
> enzymatically copying an oligonucleotide target using at least one nucleotide analogue, thereby producing multiple oligonucleotide analogue targets;
>
> selecting said oligonucleotide analogue targets such that said oligonucleotide analogue targets bind to the complementary oligonucleotide probes coupled to a solid surface at known locations of an array with a similar hybridization stability across the array;
>
> hybridizing the oligonucleotide analogue targets to complementary oligonucleotide probes; and
>
> detecting whether at least one of said oligonucleotide analogue targets binds to said complementary oligonucleotide acid probe.

The subject matter of this claim is supported by the disclosure of the Grant Application as follows. The Grant Application discloses a method for detecting nucleic acid sequence differences in target nucleic acids using an array of oligonucleotides on a solid support to capture complementary zip code tails on LDR products (Id. at 11, lines 4-6). Thymine can be replaced by 5-propynyl uridine when used within DNA or PNA address sequences (complementary zip codes) as well as in the DNA zip code sequences (Id. at 24, lines 8-9). Further, the Grant Application discloses LDR products having a "zip code" tail, which will be selectively captured by a complementary zip code (DNA oligonucleotides or PNA oligonucleotides) on a solid support (Id. at 11, lines 4-7; Id. at 19, lines 42-44) and that zip code (oligonucleotide) capture can be optimized by narrowing the thermal stability (T_m) difference between duplexes formed by LDR products having a zip code tail and the complementary zip codes (DNA or PNA oligonucleotides) on the solid support hybridized to one another resulting from differences in G•C/A•T content (Id. at 22, lines 18-20). Moreover, the Grant Application discloses raising the T_m for all zip code/address duplexes (Id. at 22, lines 18-24). The T_m for the correct PNA/DNA hybridization is at least 48 °C higher than any incorrect hybridization and neighboring 24-mers are separated by 12-mers, which do not hybridize with anything (Id. at 31, lines 23-25). In addition, the Grant Application discloses LDR products having a "zip code" tail, which hybridize to complementary zip codes on a solid support (Id. at 11, lines 5-6) and discloses detecting labels of oligonucleotide probes hybridized to the solid support (Id. at 19, line 44, to 21, line 25; Figure 2).

11. Claim 28 of the present application is directed to:

> A method of making an array of oligonucleotide probes comprising:
> providing a plurality of oligonucleotide analogue probes having at least one oligonucleotide analogue, said oligonucleotide analogue probes having different sequences at known locations on an array, and
> selecting the oligonucleotide analogue probes to hybridize with complementary oligonucleotide target sequences under hybridization conditions such that said oligonucleotide analogue probes bind to complementary oligonucleotide targets with a similar hybridization stability across the array.

The subject matter of this claim is supported by the disclosure of the Grant Application as follows. The Grant Application discloses an array of oligonucleotides on a solid support (Id.

at 19, lines 36-44) and its use in a method for detecting nucleic acid sequence differences in target nucleic acids by hybridization of nucleic acid sequences to the oligonucleotides on the array (see, e.g., Id. at 11, lines 4-6; Id. at 16, lines 1-14). DNA oligonucleotides or PNA oligomers (complementary zip codes) are attached at different locations to a solid support (Id. at 21, lines 27-30; Id. at 22, lines 9-10; Id. at 25, lines 5-8). Thymine can be replaced by 5-propynyl uridine when used within DNA or PNA address sequences (attached to a support) as well as in DNA zip code sequences (Id. at 24, lines 8-9). Further, the Grant Application discloses a plurality of DNA oligonucleotides or PNA oligomers attached to the solid support, each having different nucleotide sequences (Id. at 19, lines 36-44), PNA oligomers in unique addressable arrays (Id. at 21, lines 27-30), and a PNA oligonucleotide array with 1,296 addresses (Id. at 21, lines 53-54). Moreover, the Grant Application discloses a complementary zip code at a discrete address on a two-dimensional array (Id. at 35, lines 16-18). LDR products having a "zip code" tail will be selectively captured by a complementary zip code (DNA oligonucleotides or PNA oligonucleotides) on a solid support (Id. at 11, lines 4-7; Id. at 19, lines 42-44) and zip code (oligonucleotide) capture can be optimized by narrowing the thermal stability (T_m) difference between duplexes formed by LDR products having a zip code tail and the complementary zip codes (DNA or PNA oligonucleotides) on the solid support hybridized to one another resulting from differences in G•C/A•T content (Id. at 22, lines 18-20). Further, the Grant Application discloses raising the T_m for all zip code/address duplexes (Id. at 22, lines 18-24). Moreover, the Grant Application discloses that the T_m for the correct PNA/DNA hybridization is at least 48 °C higher than any incorrect hybridization and neighboring 24-mers are separated by 12-mers, which do not hybridize with anything (Id. at 31, lines 23-25).

12. Claim 33 of the present application is directed to:

> A composition for analyzing interactions between oligonucleotide targets and oligonucleotide probes comprising:
> a solid substrate and
> an array of a plurality of oligonucleotide analogue probes coupled to the solid substrate, wherein the oligonucleotide analogue probes have different sequences and are selected to hybridize to complementary oligonucleotide targets under uniform hybridization conditions.

The subject matter of this claim is supported by the disclosure of the Grant Application as follows. The Grant Application discloses an array of oligonucleotides on a solid support (Id.

at 19, lines 36-44) and its use in a method for detecting nucleic acid sequence differences in target nucleic acids by hybridization of nucleic acid sequences to the oligonucleotides on the array (see, e.g., Id. at 11, lines 4-6; Id. at 16, lines 1-14). A plurality of DNA oligonucleotides or peptide nucleotide analogue (PNA) oligomers (complementary zip codes) are attached at different locations to a solid support, each having different nucleotide sequences (Id. at 21, lines 27-30; Id. at 22, lines 9-10; Id. at 25, lines 5-8; Id. at 19, lines 36-44). Further, the Grant Application discloses that thymine can be replaced by 5-propynyl uridine when used within DNA or PNA address sequences (attached to a support) as well as in DNA zip code sequences (Id. at 24, lines 8-9) and discloses PNA oligomers in unique addressable arrays (Id. at 21, lines 27-30). Moreover, the Grant Application discloses a PNA oligonucleotide array with 1,296 addresses (Id. at 21, lines 53-54), a complementary zip code at a discrete address on a two-dimensional array (Id. at 35, lines 16-18), and LDR products having a "zip code" tail, which will be selectively captured by a complementary zip code (DNA oligonucleotides or PNA oligonucleotides) on a solid support (Id. at 11, lines 4-7; Id. at 19, lines 42-44). Zip code (oligonucleotide) capture can be optimized by narrowing the thermal stability (T_m) difference between duplexes formed by LDR products having a zip code tail and the complementary zip codes (DNA or PNA oligonucleotides) on the solid support hybridized to one another resulting from differences in G•C/A•T content (Id. at 22, lines 18-20). The Grant Application also discloses raising the T_m for all zip code/address duplexes (Id. at 22, lines 18-24). Further, the Grant Application discloses that the T_m for the correct PNA/DNA hybridization is at least 48 °C higher than any incorrect hybridization and neighboring 24-mers are separated by 12-mers, which do not hybridize with anything (Id. at 31, lines 23-25).

13. Claim 34 of the present application is directed to:

> A composition for analyzing interactions between oligonucleotide targets and oligonucleotide probes comprising:
> an array of a plurality of oligonucleotide probes having different sequences hybridized to complementary oligonucleotide analogue targets, wherein the oligonucleotide analogue targets hybridize to complementary oligonucleotide probes under uniform hybridization conditions.

The subject matter of this claim is supported by the disclosure of the Grant Application as follows. The Grant Application discloses an array of oligonucleotides on a solid support (Id. at 19, lines 36-44) and its use in a method for detecting nucleic acid sequence differences in

target nucleic acids by hybridization of nucleic acid sequences to the oligonucleotides on the array (see, e.g., Id. at 11, lines 4-6; Id. at 16, lines 1-14). In addition, the Grant Application discloses an array of oligonucleotides on a solid support (Id. at 19, lines 36-44), a plurality of different sequences attached at different locations (addresses) on the solid support (Id. at 19, lines 36-44), and a universal PNA oligonucleotide array with 1,296 addresses (Id. at 21, lines 53-54). LDR products having a zip code tail hybridize to complementary zip codes on a solid support where the complementary components are DNA or PNA (Id. at 11, lines 4-6) and thymine is replaced with 5-propynyl uridine when used within DNA or PNA address sequences (complementary oligonucleotide probes) as well as in the DNA zip code sequences (Id. at 11, lines 8-10; Id. at 24, lines 8-9). Moreover, the Grant Application discloses that zip code (oligonucleotide) capture can be optimized by narrowing the thermal stability (T_m) difference between duplexes formed by LDR products having a zip code tail and the complementary zip codes (DNA or PNA oligonucleotides) on the solid support hybridized to one another resulting from differences in G•C/A•T content (Id. at 22, lines 18-20). In addition, the Grant Application discloses raising the T_m for all zip code/address duplexes (Id. at 22, lines 18-24). Further, the Grant Application discloses that the T_m for the correct PNA/DNA hybridization is at least 48 °C higher than any incorrect hybridization and neighboring 24-mers are separated by 12-mers, which do not hybridize with anything (Id. at 31, lines 23-25).

14. Claim 35 of the present application is directed to:

> A method of analyzing interactions between oligonucleotide targets and oligonucleotide probes comprising:
>
> providing on a solid substrate an oligonucleotide analogue array comprising a plurality of oligonucleotide analogue probes having different sequences;
>
> exposing said oligonucleotide analogue probe array to a plurality of oligonucleotide targets under conditions effective to permit the plurality of oligonucleotide analogue probes to hybridize to complementary target oligonucleotides under uniform hybridization conditions; and
>
> determining whether an oligonucleotide analogue probe of said oligonucleotide probe array hybridizes to at least one of the oligonucleotide targets.

The subject matter of this claim is supported by the disclosure of the Grant Application as follows. The Grant Application discloses a method for detecting nucleic acid sequence

differences in target nucleic acids using an array of oligonucleotides on a solid support to capture complementary zip code tails on LDR products (Id. at 11, lines 4-6). In addition, the Grant Application discloses a plurality of DNA oligonucleotides or PNA oligomers (complementary zip codes) attached at different locations to a solid support, each having different nucleotide sequences (Id. at 21, lines 27-30; Id. at 22, lines 9-10; Id. at 25, lines 5-8; Id. at 19, lines 36-44). Thymine can be replaced by 5-propynyl uridine when used within DNA or PNA address sequences (attached to a support) as well as in DNA zip code sequences (Id. at 24, lines 8-9). Moreover, the Grant Application discloses PNA oligomers in unique addressable arrays (Id. at 21, lines 27-30), a PNA oligonucleotide array with 1,296 addresses (Id. at 21, lines 53-54), and a complementary zip code at a discrete address on a two-dimensional array (Id. at 35, lines 16-18). Further, the Grant Application discloses contacting ligated reaction product with a solid support including complementary zip codes (DNA oligonucleotides or PNA oligomers) which are complementary to addressable array-specific portions (zip codes) of a ligated product sequence (Id. at 11, lines 4-7) and discloses that the zip code tail of the ligated reaction product will be captured by a complementary zip code on the solid support (Id. at 11, lines 5-6). The nucleotide analogue, 5-propynyl uridine, can be incorporated into DNA zip code and PNA address sequences (Id. at 11, lines 8-10). The Grant Application discloses that zip code (oligonucleotide) capture can be optimized by narrowing the thermal stability (T_m) difference between duplexes formed by LDR products having a zip code tail and the complementary zip codes (DNA or PNA oligonucleotides) on the solid support hybridized to one another resulting from differences in G•C/A•T content (Id. at 22, lines 18-20). In addition, the Grant Application discloses raising the T_m for all zip code/address duplexes (Id. at 22, lines 18-24). Further, the Grant Application discloses that the T_m for the correct PNA/DNA hybridization is at least 48 °C higher than any incorrect hybridization and neighboring 24-mers are separated by 12-mers, which do not hybridize with anything (Id. at 31, lines 23-25). Ligation products are identified by detecting the labels of oligonucleotide probes hybridized to the solid support (Id. at 19, line 44, to 21, line 25; Figure 2).

15. Claim 36 of the present application is directed to:

> A method of detecting an oligonucleotide target comprising:
>
> enzymatically copying an oligonucleotide target using at least one nucleotide analogue, thereby producing multiple oligonucleotide analogue targets;

> providing on a solid substrate an oligonucleotide array comprising a plurality of oligonucleotide probes selected to hybridize to complementary oligonucleotide analogue targets under uniform hybridization conditions;
>
> exposing the oligonucleotide analogue targets to the oligonucleotide array under conditions effective to permit the oligonucleotide probes to hybridize to complementary oligonucleotide analogue targets; and
>
> detecting whether at least one of the oligonucleotide analogue targets hybridizes to a complementary oligonucleotide probe.

The subject matter of this claim is supported by the disclosure of the Grant Application as follows. The Grant Application discloses a method for detecting nucleic acid sequence differences in target nucleic acids using an array of oligonucleotides on a solid support to capture complementary zip code tails on LDR products (Id. at 11, lines 4-6). Thymine can be replaced by 5-propynyl uridine when used within DNA or PNA address sequences (complementary zip codes) as well as in the DNA zip code sequences (Id. at 24, lines 8-9). Further, the Grant Application discloses LDR products having a "zip code" tail, which will be selectively captured by a complementary zip code (DNA oligonucleotides or PNA oligonucleotides) on a solid support (Id. at 11, lines 4-7; Id. at 19, lines 42-44) and that zip code (oligonucleotide) capture can be optimized by narrowing the thermal stability (T_m) difference between duplexes formed by LDR products having a zip code tail and the complementary zip codes (DNA or PNA oligonucleotides) on the solid support hybridized to one another resulting from differences in G•C/A•T content (Id. at 22, lines 18-20). Moreover, the Grant Application discloses raising the T_m for all zip code/address duplexes (Id. at 22, lines 18-24). In addition, the Grant Application discloses that the T_m for the correct PNA/DNA hybridization is at least 48 °C higher than any incorrect hybridization and neighboring 24-mers are separated by 12-mers, which do not hybridize with anything (Id. at 31, lines 23-25). LDR products having a "zip code" tail hybridize to complementary zip codes on a solid support (Id. at 11, lines 5-6). The Grant Application discloses detecting labels of oligonucleotide probes hybridized to the solid support (Id. at 19, line 44, to 21, line 25; Figure 2).

16. Claim 37 of the present application is directed to:

> A method of making an array of oligonucleotide probes comprising;
>
> providing, on an array, a plurality of oligonucleotide analogue probes having at least one

oligonucleotide analogue and different sequences, wherein the oligonucleotide analogue probes are selected to hybridize to complementary oligonucleotide targets under uniform hybridization conditions.

The subject matter of this claim is supported by the disclosure of the Grant Application as follows. The Grant Application discloses an array of oligonucleotides on a solid support (Id. at 19, lines 36-44) and its use in a method for detecting nucleic acid sequence differences in target nucleic acids by hybridization of nucleic acid sequences to the oligonucleotides on the array (see, e.g., Id. at 11, lines 4-6; Id. at 16, lines 1-14). DNA oligonucleotides or PNA oligomers (complementary zip codes) are attached at different locations to a solid support (Id. at 21, lines 27-30; Id. at 22, lines 9-10; Id. at 25, lines 5-8) and thymine can be replaced by 5-propynyl uridine when used within DNA or PNA address sequences (attached to a support) as well as in DNA zip code sequences (Id. at 24, lines 8-9). Further, the Grant Application discloses a plurality of DNA oligonucleotides or PNA oligomers attached to the solid support, each having different nucleotide sequences (Id. at 19, lines 36-44), PNA oligomers in unique addressable arrays (Id. at 21, lines 27-30), and a PNA oligonucleotide array with 1,296 addresses (Id. at 21, lines 53-54). A complementary zip code is attached at a discrete address on a two-dimensional array (Id. at 35, lines 16-18). In addition, the Grant Application discloses LDR products having a "zip code" tail, which will be selectively captured by a complementary zip code (DNA oligonucleotides or PNA oligonucleotides) on a solid support (Id. at 11, lines 4-7; page 19, lines 42-44) and that zip code (oligonucleotide) capture can be optimized by narrowing the thermal stability (T_m) difference between duplexes formed by LDR products having a zip code tail and the complementary zip codes (DNA or PNA oligonucleotides) on the solid support hybridized to one another resulting from differences in G•C/A•T content (Id. at 22, lines 18-20). Further, the Grant Application discloses raising the T_m for all zip code/address duplexes (Id. at 22, lines 18-24). Moreover, the Grant Application discloses that the T_m for the correct PNA/DNA hybridization is at least 48 °C higher than any incorrect hybridization and neighboring 24-mers are separated by 12-mers, which do not hybridize with anything (Id. at 31, lines 23-25).

17. On May 31, 1994 to June 2, 1994, the NCI Site Special Review Subcommittee for the Grant Application visited Cornell University Medical College, New York, New York to meet with the scientists who submitted the Grant Application, including me, Donald Bergstrom, and co-inventors George Barany and Robert Hammer. As indicated in the follow-up July 20-22, 1994, Draft Review Report (attached hereto at Appendix 2)

regarding the Grant Application, the NCI Site Special Review Subcommittee included Steven P.A. Fodor, Ph.D., Scientific Director and Chief Technical Officer of Affymetrix, Inc., Santa Clara, California (Appendix 2 at p. 32), who I remember was in attendance at the site visit. Dr. Fodor's presence at the site visit (as well as on the review subcommittee) caused a great deal of concern amongst me and the other grant applicants. Having served on a number of NCI site review subcommittees, it is my understanding that subcommittee members should not review grant applications where they have a conflict of interest or the appearance of one. For example, where a subcommittee member has a financial interest in work closely related or competing with the subject matter of the grant application, a conflict of interest arises and that subcommittee member should not participate in deliberations and actions on the grant application. Since it was well known at the time of the site visit that Affymetrix, Inc. was developing oligonucleotide arrays for use in detecting single base mutations, I and other grant applicants wondered why Dr. Fodor did not remove himself from the review subcommittee.

18. Prior to the site visit I expressed my concern about Dr. Fodor's presence at the site visit to NCI; however, with assurance by NCI that it would guard against conflicts of interest problems, Dr. Fodor remained on the review subcommittee. I continued to be concerned about Dr. Fodor's presence on the subcommittee and, after learning of work subsequently presented by Affymetrix scientists, sent a letter with Donald Bergstrom on November 11, 1994, to Dr. James Jacobson of NCI (attached hereto at Appendix 3), in which we again protested the presence of Dr. Fodor on the site review subcommittee on the grounds that he may have had a conflict of interest.

19. My co-inventors and I regard the subject matter of Project 5 of the Grant Application as our invention and have fully disclosed it in the above patent application. In particular, a method of detecting nucleic acid sequence differences in target nucleic acids by the use of a solid support with an array of DNA oligonucleotides or PNA oligomers to capture and detect the products of LDR is described on page 1, lines 13-22, page 8, lines 10-13, page 30, lines 32-34, and page 40, lines 35-37 of our above-identified patent application. Our patent application further teaches (at page 16, line 23 to page 19, line 5 and in the drawings that these pages describe) that each ligation product of the LDR process is provided with an addressable array-specific portion which is selectively captured by a complementary capture probe at a particular location on the solid support. In fact, the same figures discussed in preceding paragraph 6 are Figures 13A-C of the present application which are described in the paragraph bridging pages 25 and 26 of the present application. The present application further discloses that the array capture probes on the solid support can be designed to achieve

optimal hybridization to the addressable array-specific portion of the ligation product (see page 35, lines 12-15 and page 49, lines 16-18). This is achieved by increasing thermal stability through reduction of the difference in melting temperature (i.e. T_m) between the different duplexes formed by LDR products having an addressable array-specific portion and the complementary capture probe (i.e. DNA oligonucleotides or PNA oligomers) hybridized to one another on the solid support (Id. at page 35, lines 12-15). These differences in melting temperature result from differences in G•C/A•T content (Id.).

20. I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under section 1001 of Title 18 of the United States Code, and that such willful false statements may jeopardize the validity of the application or any patent issuing thereon.

Date: 12/17/01

Dr. Francis Barany

Francis Barany

Appendix 1

(To Declaration of Francis Barany under 37 CFR § 1.608(b))

AA

Form Approved Through 6.30.94
OMB No. 0925-0001

DEPARTMENT OF HEALTH AND HUMAN SERVICES
PROGRAM PUBLIC HEALTH SERVICE
PROJECT **GRANT APPLICATION**
Follow instructions carefully. Type in the unshaded areas only. Type density must be 10 c.p.i.

LEAVE BLANK FOR PHS USE ONLY.		
Type	Activity	Number
Review Group		Formerly
Council/Board *(Month, Year)*		Date Received

1. TITLE OF PROJECT (Do not exceed 56 typewriter spaces.)
NEW METHODS FOR CANCER DETECTION

2a. RESPONSE TO SPECIFIC REQUEST FOR APPLICATIONS OR PROGRAM ANNOUNCEMENT ☒ NO ☐ YES *(If "YES," state number and title)*
Number: Title: CONFIDENTIAL INFORMATION

2b. TYPE OF GRANT PROGRAM P01

3. PRINCIPAL INVESTIGATOR/PROGRAM DIRECTOR

3a. NAME *(Last, first, middle)*
BARANY, FRANCIS

3b. DEGREE(S)
PH.D.

3c. SOCIAL SECURITY NO.
REDACTED

3d. POSITION TITLE
ASSOCIATE PROFESSOR

3e. MAILING ADDRESS *(Street, city, state, zip code)*
CORNELL UNIV. MEDICAL COLLEGE
1300 YORK AVENUE
NEW YORK, NY 10021

3f. DEPARTMENT, SERVICE, LABORATORY, OR EQUIVALENT
MICROBIOLOGY

3g. MAJOR SUBDIVISION
CORNELL UNIV. MEDICAL COLLEGE

3h. TELEPHONE AND FAX *(Area code, number and extension)*
TEL: 212 746-6509
FAX: 212 746-8587

BITNET/INTERNET ADDRESS BARANY@CUMC.CORNELL.EDU

4. HUMAN SUBJECTS — If "Yes," exemption no. or IRB approval date — 4b. Assurance of compliance no.
☒ 4a. NO ☐ YES

5. VERTEBRATE ANIMALS — If "Yes," IACUC approval date — 5b. Animal welfare assurance no.
☒ 5a. NO ☐ YES

6. DATES OF ENTIRE PROPOSED PROJECT PERIOD		7. COSTS REQUESTED FOR INITIAL BUDGET PERIOD		8. COSTS REQUESTED FOR ENTIRE PROPOSED PROJECT PERIOD	
From (MMDDYY)	Through (MMDDYY)	7a. Direct Costs ($)	7b. Total Costs ($)	8a. Direct Costs ($)	8b. Total Costs ($)
120194	113099	$971,041	$1,224,862	$5,295,869	$6,995,293

9. PERFORMANCE SITES *(Organizations and addresses)*
CORNELL UNIV. MEDICAL COLLEGE
1300 YORK AVENUE
NEW YORK, NY 10021

10. INVENTIONS AND PATENTS *(Competing continuation application only)*
☐ NO ☐ YES If "YES," ☐ Previously reported ☐ Not previously reported

11. NAME OF APPLICANT ORGANIZATION
CORNELL UNIVERSITY MEDICAL COLLEGE
ADDRESS 1300 YORK AVENUE
NEW YORK, NY 10021

12. TYPE OF ORGANIZATION
☐ Public: *Specify* ☐ *Federal* ☐ *State* ☐ *Local*
☒ Private Nonprofit
☐ Forprofit *(General)* ☐ Forprofit *(Small Business)*

13. ENTITY IDENTIFICATION NUMBER
1131623978A1

Congressional District
14

14. BIOMEDICAL RESEARCH SUPPORT GRANT CREDIT
Code: 01 Identification: SCHOOL OF MEDICINE

15. NAME OF ADMINISTRATIVE OFFICIAL TO BE NOTIFIED IF AWARD IS MADE
PHILIP V. GIUCA
TELEPHONE 212 746-6036
FAX 212 746-8745
TITLE SENIOR ASSOCIATE DEAN
ADDRESS CORNELL UNIV. MEDICAL COLLEGE
1300 YORK AVENUE
NEW YORK, NY 10021

BITNET/INTERNET ADDRESS QMCUMC.MAIL.CORNELL.EDU

16. NAME OF OFFICIAL SIGNING FOR APPLICANT ORGANIZATION
GREGORY W. SISKIND
TELEPHONE 212 746-6020
FAX 212 746-8745
TITLE ASSOCIATE DEAN
ADDRESS CORNELL UNIV. MEDICAL COLLEGE
1300 YORK AVENUE
NEW YORK, NY 10021

BITNET/INTERNET ADDRESS QMCUMC.MAIL.CORNELL.EDU

17. PRINCIPAL INVESTIGATOR/PROGRAM DIRECTOR ASSURANCE: I agree to accept responsibility for the scientific conduct of the project and to provide the required progress reports if a grant is awarded as a result of this application. Willful provision of false information is a criminal offense (U.S. Code, Title 18, Section 1001). I am aware that any false, fictitious, or fraudulent statement may, in addition to other remedies available to the Government, subject me to civil penalties under the Program Fraud Civil Remedies Act of 1986 (45 CFR 79).

SIGNATURE OF PERSON NAMED IN 3a. *(In ink. "Per" signature not acceptable.)* Dr F. Barany — DATE 2/4/94

18. CERTIFICATION AND ACCEPTANCE: I certify that the statements herein are true and complete to the best of my knowledge, and accept the obligation to comply with Public Health Service terms and conditions if a grant is awarded as the result of this application. A willfully false certification is a criminal offense (U.S. Code, Title 18, Section 1001). I am aware that any false, fictitious, or fraudulent statement may, in addition to other remedies available to the Government, subject me to civil penalties under the Program Fraud Civil Remedies Act of 1986 (45 CFR 79).

SIGNATURE OF PERSON NAMED IN 16. *(In ink. "Per" signature not acceptable.)* Gregory W. Siskind — DATE 2/4/94

DESCRIPTION: State the application's broad, long-term objectives and specific aims, making reference to the health relatedness of the project. Describe concisely the research design and methods for achieving these goals. Avoid summaries of past accomplishments and the use of the first person. This abstract is meant to serve as a succinct and accurate description of the proposed work when separated from the application. DO NOT EXCEED THE SPACE PROVIDED.

The long range objective of this proposal is to develop sensitive and specific approaches to the detection and simultaneous identification of cancer-related, genetic alterations. Mutations and genetic aberrations have been implicated, at various steps, in the etiology and biology of tumors. Inherited mutations account for the predisposition to cancer in some families. Somatic mutations in tumor suppressor genes, oncogene amplification and viral DNA sequences have been found in cancers as well. However, the clinical use of these discoveries and research into their clinical significance has been slowed by the laborious processes by which they are detected. To apply these discoveries and explore the interactions of multiple genetic alterations, we urgently need a new technology, which is capable of being automated and has the power to detect any of a vast number of mutations.

In response to the urgent need for new methods of mutation detection, we have assembled a team of investigators whose expertise will be directed toward innovative solutions to this problem. The collaborative nature of the scientific and organizational infrastructure will facilitate the attainment of the projects' specific aims and objectives.

The specific aims of the five projects in this proposal are to: (i) develop a multiplex polymerase chain reaction/ligase detection reaction (PCR/LDR) system for the detection of inherited mutations in germline DNA and somatic mutations in tumors; (ii) develop a ligase detection reaction/ polymerase chain reaction (LDR/PCR) system for detecting gene amplifications and deletions in tumors; (iii) develop a PCR/restriction enzyme/LDR (PCR/RE/LDR) system for detecting and identifying mutations in rare cancer cells at a sensitivity of 1 in 10^6 or 1 in 10^7 by removing normal DNA sequences and selectively amplifying cancer mutations; (iv) design and synthesize nucleotide analogues for converting specific DNA sequences into restriction endonuclease recognition sites for PCR/RE/LDR mutation detection; (v) engineer a thermostable ligase with greater fidelity to enhance LDR and LCR specificity; (vi) design and synthesis oligonucleotide or peptide nucleic acid (PNA) addressable arrays for the simultaneous detection of multiplex LDR and LCR products; and (vii) explore the ability of these technologies to further our understanding and clinical management of lung, colon, breast and cervical cancers.

PERSONNEL ENGAGED ON PROJECT, INCLUDING CONSULTANTS/COLLABORATORS. *Use continuation pages as needed* to provide the required information in the format shown below on *all* individuals participating in the project.

Name AGGARWAL, Aneel	Degree(s) Ph.D.	Social Security # REDACTED
Position Title Assistant Professor	D.O.B. REDACTED	Role on Project Co-investigator
Organization College of Physicians & Surgeons of Columbia University		Department Biochem & Biophys
Name AHNEN, Dennis	Degree(s) M.D.	Social Security # REDACTED
Position Title VA Clinical Investigator /Associate Professor	D.O.B. REDACTED	Role on Project Collaborator
Organization Veterans Affairs Medical Center & Univ. Colorado School Of Medicine		Department Medicine
Name BARANY, Francis	Degree(s) Ph.D.	Social Security # REDACTED
Position Title Associate Professor	D.O.B. REDACTED	Role on Project Prin. Investig.
Organization Cornell Univeristy Medical College		Department Microbiology
Name BARANY, George	Degree(s) Ph. D.	Social Security # REDACTED
Position Title Professor	D.O.B. REDACTED	Role on Project Co-investigator
Organization University of Minnesota		Department Chemistry
Name BATT,Carl	Degree(s) Ph.D.	Social Security # REDACTED
Position Title Associate Professor	D.O.B. REDACTED	Role on Project Collaborator
Organization Cornell University		Department Food Science
Name BERGSTROM, Donald	Degree(s) Ph.D.	Social Security # REDACTED
Position Title Professor	D.O.B. REDACTED	Role on Project Co-investigator
Organization Purdue Univ. School of Pharmacy &Pharm. Sciences		Department Medicinal Chem.
Name BUNK, Michael	Degree(s) Ph.D.	Social Security # REDACTED
Position Title Director, Research Management	D.O.B. REDACTED	Role on Project Administrator
Organization Strang Cancer Prevention Center		Department Res. Management

Name	COOK, Ronald	Degree(s)	Ph. D.	Social Security #	REDACTED
Position Title	President/Chief Technical Officer	D.O.B.	REDACTED	Role on Project	Collaborator
Organization	Siris Laboratories			Department	Chemistry
Name	COTHERN, Melissa	Degree(s)	B.S.	Social Security #	REDACTED
Position Title	Graduate Student	D.O.B.	REDACTED	Role on Project	Synthetic Chem.
Organization	Louisiana State University			Department	Chemistry
Name	COULL, James	Degree(s)	Ph.D.	Social Security #	REDACTED
Position Title	Group Manager of Specialty Chemistry	D.O.B.	REDACTED	Role on Project	Collaborator
Organization	Millipore Corporation			Department	Specialty Chemistry
Name	COURVALIN, Patrice	Degree(s)	M. D.	Social Security #	not applicable
Position Title	Professor, Associate Chairman	D.O.B.	REDACTED	Role on Project	Collaborator
Organization	Institut Pasteur			Department	Bacteriol. & Mycol.
Name	DAY, Darren	Degree(s)	Ph.D.	Social Security #	REDACTED
Position Title	Research Associate	D.O.B.	REDACTED	Role on Project	
Organization	Cornell Univeristy Medical College			Department	Microbiology
Name	FISHMAN, Jack	Degree(s)	Ph. D.	Social Security #	REDACTED
Position Title	Director of Research	D.O.B.	REDACTED	Role on Project	Collaborator
Organization	Strang Cancer Research Lab			Department	Horm. Carcinogens
Name	FRANKLIN, Wilbur	Degree(s)	M.D.	Social Security #	REDACTED
Position Title	Professor, Director of Tissue Bank	D.O.B.	REDACTED	Role on Project	Collaborator
Organization	University of Colorado, School of Medicine			Department	Pathology
Name	FRIEND, Steven	Degree(s)	M.D./Ph.D.	Social Security #	REDACTED
Position Title	Assistant Professor	D.O.B.	REDACTED	Role on Project	Collaborator
Organization	Harvard Medical School, MGH Cancer Center			Department	Cell & Dev. Biol.
Name	GELFAND, David	Degree(s)	Ph.D.	Social Security #	REDACTED
Position Title	Director , Core Technology	D.O.B.	REDACTED	Role on Project	Collaborator
Organization	Roche Molecular Systems			Department	Protein Core Res.
Name	GILES, Aaron	Degree(s)	B.S.	Social Security #	REDACTED
Position Title	Computer Programmer /Analyst II	D.O.B.	REDACTED	Role on Project	
Organization	Cornell University Medical College			Department	Acad. Computing
Name	GROSSMAN, Larry	Degree(s)	Ph.D.	Social Security #	REDACTED
Position Title	Professor, Associate Chairman	D.O.B.	REDACTED	Role on Project	Collaborator
Organization	Wayne State University, School of Medicine			Department	Mol. Biol. & Gen.
Name	HACKETT, Neil	Degree(s)	Ph.D.	Social Security #	REDACTED
Position Title	Assistant Professor	D.O.B.	REDACTED	Role on Project	Co-investigator
Organization	Cornell Univeristy Medical College			Department	Microbiology
Name	HAMMER, Robert	Degree(s)	Ph.D.	Social Security #	REDACTED
Position Title	Assistant Professor	D.O.B.	REDACTED	Role on Project	Co-investigator
Organization	Louisiana State University			Department	Chemistry
Name	HERRARA, Vicky	Degree(s)	M.D.	Social Security #	REDACTED
Position Title	Assistant Professor	D.O.B.	REDACTED	Role on Project	Collaborator
Organization	Boston University School of Medicine			Department	Medicine
Name	HOFFMAN, Eric	Degree(s)	Ph.D.	Social Security #	REDACTED
Position Title	Assistant Professor	D.O.B.	REDACTED	Role on Project	Collaborator
Organization	University of Pittsburgh, School of Medicine			Department	Mol. Gen. Biochem
Name	KENNEDY, Timothy	Degree(s)	M.D.	Social Security #	REDACTED
Position Title	Medical Director	D.O.B.	REDACTED	Role on Project	Collaborator
Organization	Lung Cancer Institut of Colorado			Department	Institute
Name	KEW, Olen	Degree(s)	Ph.D.	Social Security #	REDACTED
Position Title	Chief of Molecular Virology Section	D.O.B.	REDACTED	Role on Project	Collaborator
Organization	National Center of Infectious Diseases, CDC			Department	Viral Division
Name	KOLLER, Antje	Degree(s)	B.A. equiv.	Social Security #	REDACTED
Position Title	Technician	D.O.B.	REDACTED	Role on Project	
Organization	Cornell Univeristy Medical College			Department	Microbiology
Name	KOVACH, John	Degree(s)	M.D.	Social Security #	REDACTED
Position Title	Professor, Chairman, Director NCI Cancer Ctr.	D.O.B.	REDACTED	Role on Project	Collaborator
Organization	Mayo Clinic			Department	Oncology

Name LU, Jing	Degree(s) B.A.	Social Security # REDACTED	
Position Title Technician	D.O.B. REDACTED	Role on Project	
Organization Cornell Univeristy Medical College		Department Microbiology	
Name LUBIN, Matthew	Degree(s) M.D.	Social Security # REDACTED	
Position Title Director of Medical Genetics	D.O.B. REDACTED	Role on Project Co-investigator	
Organization Strang Cancer Prevention Center		Department Medical Genetics	
Name LUO, Jianying	Degree(s) Ph.D.	Social Security # REDACTED	
Position Title Research Associate	D.O.B. REDACTED	Role on Project	
Organization Cornell Univeristy Medical College		Department Microbiology	
Name MILLER, Gary	Degree(s) M.D./Ph.D.	Social Security # REDACTED	
Position Title Associate Professor, Director, Histo. Pathology	D.O.B.	Role on Project Collaborator	
Organization Tissue Procurement Core Lad, Univ. of Colorado School of Medicine		Department Pathology	
Name NORTHRUP, Allen	Degree(s) Ph.D.	Social Security # REDACTED	
Position Title Principal Engineer / Adjunct Assistant Prof.	D.O.B. REDACTED	Role on Project Collaborator	
Organization Lawrence Livermore National Lab / U. C. S. F. Medical Center		Department Radiology	
Name OSBORNE, Michael	Degree(s) M.D.	Social Security # REDACTED	
Position Title Director	D.O.B. REDACTED	Role on Project Collaborator	
Organization Strang Cancer Prevention Center		Department Surgery	
Name PERSING, David	Degree(s) M.D./Ph.D.	Social Security # REDACTED	
Position Title Senior Associate Consultant / Assistant Prof.	D.O.B. REDACTED	Role on Project Collaborator	
Organization Mayo Foundation		Department Lab. Med.& Pathol.	
Name PROUDFOOT, Susan	Degree(s) M.S.H.A.	Social Security # REDACTED	
Position Title Executive Director	D.O.B. REDACTED	Role on Project Collaborator	
Organization Lung Cancer Institute of Colorado		Department Lung Institute	
Name REZNIKOV, Leonid	Degree(s) M.D./Ph.D.	Social Security # REDACTED	
Position Title Research Fellow	D.O.B. REDACTED	Role on Project	
Organization The Children's Hospital & Univ. Colorado Health Sciences Center		Department Pathology	
Name RIGAS, Basil	Degree(s) M.D.	Social Security # REDACTED	
Position Title Associate Professor	D.O.B. REDACTED	Role on Project Collaborator	
Organization Cornell University Medical College		Department Microbiol.& Med.	
Name ROBERTS, Richard	Degree(s) Ph.D.	Social Security # REDACTED	
Position Title Director of Research	D.O.B. REDACTED	Role on Project Collaborator	
Organization New England BioLabs, Inc.		Department Eucaryotic Biol.	
Name SHILDKRAUT, Ira	Degree(s) Ph.D.	Social Security # REDACTED	
Position Title Director of Research	D.O.B.	Role on Project Collaborator	
Organization New England BioLabs, Inc.		Department Research & Devel.	
Name SILVERSTEIN, Saul	Degree(s) Ph.D.	Social Security # REDACTED	
Position Title Professor	D.O.B. REDACTED	Role on Project Collaborator	
Organization Columbia University		Department Microbiology	
Name SNINSKY, John	Degree(s) Ph.D.	Social Security # REDACTED	
Position Title Senior Director of Research	D.O.B. REDACTED	Role on Project Collaborator	
Organization Roche Molecular Systems, Inc.		Department PCR Research	
Name SOBEL, Mark	Degree(s) M.D./Ph.D.	Social Security # REDACTED	
Position Title Chief, Molecular Pathology Section	D.O.B. REDACTED	Role on Project Collaborator	
Organization National Cancer Institute		Department Pathology	
Name SOMMER, Steven	Degree(s) M.D./Ph.D.	Social Security # REDACTED	
Position Title Associate Professor	D.O.B. REDACTED	Role on Project Collaborator	
Organization Mayo Clinic		Department Biochem, Mol.Biol.	
Name SOUSSI, Thierry	Degree(s) Ph.D.	Social Security # Not applicable	
Position Title Professor	D.O.B. REDACTED	Role on Project Collaborator	
Organization Institute de Genetique Moleculaire		Department Molecular Genetics	
Name SWERDLOW, Harold	Degree(s) Ph.D.	Social Security # REDACTED	
Position Title Research Associate	D.O.B. REDACTED	Role on Project Collaborator	
Organization University of Utah		Department Human Genetics	
Name VAGNER, Josef	Degree(s) Ph.D.	Social Security # REDACTED	
Position Title Postdoctoral Fellow	D.O.B. REDACTED	Role on Project	
Organization University of Minnesota		Department Chemistry	

Name	VAGNEROVA, Lydie	Degree(s)	B.S.	Social Security #	Pending
Position Title	Research Technician	D.O.B.	REDACTED	Role on Project	
Organization	University of Minnesota			Department	Chemistry
Name	WANG, Guangyi	Degree(s)	Ph.D.	Social Security #	REDACTED
Position Title	Postdoctoral Fellow	D.O.B.	REDACTED	Role on Project	
Organization	Purdue University			Department	Medicinal Chem.
Name	WHITE, Perrin	Degree(s)	M.D./Ph.D.	Social Security #	REDACTED
Position Title	Professor	D.O.B.	REDACTED	Role on Project	Collaborator
Organization	Cornell University Medical College			Department	Pediatrics
Name	WILSON, Geoffrey	Degree(s)	Ph.D.	Social Security #	REDACTED
Position Title	Research Group Leader	D.O.B.	REDACTED	Role on Project	Collaborator
Organization	New England Biolabs, Inc.			Department	Research
Name	WILSON, Vincent	Degree(s)	Ph.D.	Social Security #	REDACTED
Position Title	Associate Professor / Director	D.O.B.	REDACTED	Role on Project	Co-investigator
Organization	The Children's Hospital & Univ. of Colorado School Of Medicine			Department	Pathology
Name	WINN-DEEN, Emily	Degree(s)	Ph.D.	Social Security #	REDACTED
Position Title	Staff Scientist	D.O.B.	REDACTED	Role on Project	Collaborator
Organization	Applied Biosystems, Inc.			Department	Research & Devel.
Name	ZHANG, Peiming	Degree(s)	Ph.D.	Social Security #	REDACTED
Position Title	Postdoctoral Research Assistant	D.O.B.	REDACTED	Role on Project	
Organization	Purdue University			Department	Medicinal Chem.

TABLE OF PROPORTIONAL EFFORT OF INVESTIAGTORS

Investigator	Project 1	Project 2	Project 3	Project 4	Project 5	Core A	Core B	Core C	Total	Other Support
A. Aggarwal				10%					10%	75%
F. Barany		15%		10%			10%	5%	40%	65%
G. Barany					10%				10%	67%
D. Bergstrom			10%						10%	50%
N. Hackett						20%			20%	50%
R. Hammer			26%						26%	35%
M. Lubin		10%					10%		20%	0%
V. Wilson	20%								20%	50%

TABLE OF CONTENTS

Project 2. Genetic Markers of Breast and Cervical Cancer
Project Leader: Dr. F. Barany
Project Co-Leader: Dr. M. Lubin

Project 3. Design and Synthesis of Nucleotide Analogues
Project Leader: Dr. D. Bergstrom
Project Co-Leader: Dr. R. Hammer

Project 4. Engineering an Improved Thermostable Ligase
Project Leader: Dr. F. Barany
Project Co-Leader: Dr. A. Aggarwal

Project 5. Design and Synthesis of DNA and PNA Arrays

Project Leader: Dr. G. Barany

Core A. Informatic Support for Cancer Detection Methods

Core Leader: Dr. N. Hackett

Core B. Instrumentation and Mutation Detection

Core Leader: Dr. F. Barany

Core Co-Leader: Dr. M. Lubin

Core C. Administrative Core

Core Leader: Dr. F. Barany

Core Co-Leader: Dr. M. Bunk

Project 5.

Design and Synthesis of DNA and PNA Arrays

Project Leader: George Barany
University of Minnesota

DESCRIPTION: State the application's broad, long-term objectives and specific aims, making reference to the health relatedness of the project. Describe concisely the research design and methods for achieving these goals. Avoid summaries of past accomplishments and the use of the first person. This abstract is meant to serve as a succinct and accurate description of the proposed work when separated from the application. **DO NOT EXCEED THE SPACE PROVIDED.**

The goal of this program project is to develop methods for identifying multiple gene mutations in cancers. For maximum utility, these methods must be able to recognize and discriminate between dozens or hundreds of mutations.

To accomplish this, we propose to capture specific ligase detection reaction (LDR) products on a spatially addressable array, such that the position of a signal identifies a mutation. Each LDR product will have a "zip code" tail, which will be selectively captured by a "complementary zip code" on a solid support. The complementary componenets can be DNA oligonucleotides or peptide nucleotide analogues (PNA). PNA/DNA hybrids have significantly higher T_m values than DNA/DNA hybrids. Incorporation of the nucleotide analogue, 5-propynyluridine, into DNA zip code and PNA address sequences will further increase and optimize T_m values (Project 3). Unreacted LDR primer may therefore be washed away at high temperatures allowing for a higher sensitivity in detecting LDR products. A reusable, universal addressable array could be used for detecting a wide range of cancer mutations, genetic diseases and infectious agents.

Implementation of these concepts, with the ultimate goal of achieving reliable and efficient materials and procedures that can be incorporated into easy-to-use, automated, low-cost diagnostic devices, will follow these aims: (i) Development and evaluation of solid support materials compatible with chemical synthesis of DNA oligonucleotides and PNA oligomers, and compatible with subsequent hybridization reactions. Surfaces, beads, or membranes will be functionalized, and extended as needed with hydrophilic spacers such as heterobifunctional polyethylene glycol (PEG) and/or carbohydrates. Chemistry for linking oligomers to the solid support, and/or solid-phase assembly of oligomers, will be developed. (ii) Establishment of methodology for synthesis of spatially addressable arrays of DNA oligonucleotides and PNA oligomers. Appropriate masking technology will expose defined regions of the solid support for attachment of pre-formed oligomers, or for chain elongation to assemble the needed oligomers. In the latter mode, segment condensation will be used when possible in order to provide efficient convergent synthesis, and because chemical "failures" will become "invisible" during the subsequent hybridization. (iii) Demonstration of scope and limitations of zip code concepts. As aims (i) and (ii) come to fruition, testing will be carried out (Core B). Design of primer and zip code structures will be facilitated by the informatics collaboration (Core A).

PERSONNEL ENGAGED ON PROJECT, INCLUDING CONSULTANTS/COLLABORATORS. *Use continuation pages as needed* to provide the required information in the format shown below on *all* individuals participating in the project.

Name	BARANY, George	Degree(s)	Ph.D.	Social Security #	REDACTED
Position Title	Professor	D.O.B.	REDACTED	Role on Project	Prin. Investig.
Organization	University of Minnesota			Department	Chemistry
Name	HAMMER, Robert	Degree(s)	Ph.D.	Social Security #	REDACTED
Position Title	Assistant Professor	D.O.B.	REDACTED	Role on Project	Co-investigator
Organization	Louisiana State University			Department	Chemistry
Name	VÁGNER, Josef	Degree(s)	Ph.D.	Social Security #	REDACTED
Position Title	Postdoctoral Associate	D.O.B.	REDACTED	Role on Project	
Organization	University of Minnesota			Department	Chemistry
Name	VÁGNEROVA, Lydie	Degree(s)	B.S.	Social Security #	pending
Position Title	Research Technician	D.O.B.	REDACTED	Role on Project	
Organization	University of Minnesota			Department	Chemistry
Name		Degree(s)		Social Security #	
Position Title		D.O.B.		Role on Project	
Organization				Department	
Name		Degree(s)		Social Security #	
Position Title		D.O.B.		Role on Project	
Organization				Department	

DD

Principal Investigator/Program Director (Last, first, middle): F. BARANY, Ph.D.

DETAILED BUDGET FOR INITIAL BUDGET PERIOD
DIRECT COSTS ONLY

FROM 94/12/01 THROUGH 95/11/30

PERSONNEL (Applicant Organization Only) NAME	ROLE ON PROJECT	TYPE APPT. (months)	% EFFORT ON PROJ.	INST. BASE SALARY	DOLLAR AMOUNT REQUESTED (omit cents) SALARY REQUESTED	FRINGE BENEFITS	TOTALS
George Barany (AY)	Principal Investigator	9	5				
George Barany (SS)	Principal Investigator	3	16.7				
Josef Vagner	Post-Doc Associate	12	100				
Lydie Vagnerova	Research Technician	12	100				
PROJECT 5							
SUBTOTALS →					$55,623	$8,446	$64,069

CONSULTANT COSTS

EQUIPMENT (Itemize)

$0

SUPPLIES (Itemize by category)
Chromatography $3,000
Chemicals $5,000
Special Solvents & Reag for PNA synthesis $7,500

$15,500

TRAVEL
One trip per year for P.I. to present results $1,200

$1,200

PATIENT CARE COSTS
INPATIENT $0
OUTPATIENT $0

ALTERATIONS AND RENOVATIONS (Itemize by category)

$0

OTHER EXPENSES (Itemize by category)
See following page $5,000

$5,000

SUBTOTAL DIRECT COSTS FOR INITIAL BUDGET PERIOD **$85,769**

CONSORTIUM/CONTRACTUAL COSTS
DIRECT COSTS
INDIRECT COSTS 40% Direct

TOTAL → $34,308

TOTAL DIRECT COSTS FOR INITIAL BUDGET PERIOD (Item 7a, Face Page) → **$120,077**

PHS 398 (Rev 9/91) (Form Page 4) Page 372 DD
Number pages consecutively at the bottom throughout the application. Do not use suffixes such

BUDGET FOR ENTIRE PROPOSED PROJECT PERIOD — PROJECT 5
DIRECT COSTS ONLY

BUDGET CATEGORY TOTALS		INITIAL BUDGET PERIOD *(from page 4)*	ADDITIONAL YEARS OF SUPPORT REQUESTED 2nd	3rd	4th	5th
PERSONNEL: *Salary & fringe benefits Applicant organization only*		$64,069	$66,632	$69,297	$72,069	$74,952
CONSULTANT COSTS		$0	$0	$0	$0	$0
EQUIPMENT		$0	$2,000	$2,000	$2,000	$2,000
SUPPLIES		$15,500	$16,120	$16,765	$17,436	$18,133
TRAVEL		$1,200	$1,248	$1,298	$1,350	$1,404
PATIENT CARE COSTS	INPATIENT	$0	$0	$0	$0	$0
	OUTPATIENT	$0	$0	$0	$0	$0
ALTERATIONS AND RENOVATIONS		$0	$0	$0	$0	$0
OTHER EXPENSES		$5,000	$5,200	$5,408	$5,624	$5,849
SUBTOTAL DIRECT COSTS		$85,769	$91,200	$94,768	$98,479	$102,338
CONSORTIUM/ CONTRACTUAL COSTS		$34,308	$35,680	$37,107	$38,591	$40,135
TOTAL DIRECT COSTS		$120,077	$126,880	$131,875	$137,070	$142,473
TOTAL DIRECT COSTS FOR ENTIRE PROPOSED PROJECT PERIOD					***(Item 8a)->***	**$658,375**

JUSTIFICATION (Use continuation pages if necessary):

From Budget for Initial Period: Describe the specific functions of the personnel, collaborators, and consultants and identify individuals with appointments that are less than full time for a specific period of the year, including VA appointments.

For All Years: Explain and justify purchase of major equipment, unusual supplies requests, patient care costs, alterations and renovations, tuition remission, and donor/volunteer costs.

From Budget for Entire Period: Identify with an asterisk (*) on this page and justify any significant increase or decrease in any category over the initial budget period. Describe any change in effort of personnel.

For Competing Continuation Applications: Justify any significant increases or decreases in any category over the current level of support.

INITIAL BUDGET PERIOD:

General: This budget covers only the expenses in Dr. George Barany's laboratory at the University of Minnesota, and represents modest levels by comparison to other ongoing and past grants from NIH. The significant intellectual and experimental contributions to the project of Dr. Robert P. Hammer from Louisiana State University, and Dr. Francis Barany of Cornell University Medical College, are supported by separate budgets in this program project.

Personnel: The preparation, manipulation, and characterization of oligonucleotide and PNA building blocks and oligomers, as well as of a range of solid supports needed for synthesis and hybridizations, is quite labor-intensive and requires experienced and well-trained co-workers. Dr. Josef Vágner is a highly qualified peptide chemist who has been working in Dr. Barany's laboratory for over a

year on other projects. Ms. Vágnerova holds a degree in biochemistry and has recently arrived in this country to join the research group as a technician. These two individuals will be able to make an immediate impact to this Research Plan. Salaries follow University of Minnesota guidelines, and fringe benefits are calculated at 23.7% academic, 3.6% postdoctoral, and 27.5% civil service. Professors have 9-month academic appointments, and need to cover the 3-month summer salary from grants.

Equipment: Dr. Barany's laboratory has all of the major instrumentation required to carry out this research, with several new instruments acquired recently to replace and/or augment older models (listed with "Resources and Environment").

Supplies and Other Expenses: We are perpetually underfunded in these categories, and require a combination of grants to cover these costs. "Supplies" include chemical reagents, consumable supplies, and chromatography expenses. "Other Expenses" cover analytical fees (NMR, mass spectrometry, elemental analysis), instrument maintenance (service contracts on peptide synthesizers and amino acid analyzer shared with other grants), publication costs, communications, etc.

Travel: Funds are requested to allow attendance at one professional meeting per year in order to present results and learn of advances in scientific fields related to this proposal.

Consortium/Contractual Costs: The University of Minnesota negotiated (May 13, 1992) indirect cost rate is 40% of modified direct costs (total direct costs - equipment - graduate student benefits).

CONTINUATION YEARS:

Personnel: Dr. Barany's research program attracts a constant stream of postdoctoral applicants from good laboratories throughout the world, so there will be no difficulty in appointing individuals to continue the work after Dr. Vágner leaves. Similarly there is a good pool of candidates for technician positions. The "Research Plan" will require constant staffing at the level of two individuals.

Equipment: We request $2,000 per year to cover relatively small items of lab hardware.

Increases: Following NIH guidelines, the percentage recurring annual increase in all costs are calculated at 4%.

Five years of support are requested in order to allow enough time to show significant progress on the goals of the "Research Plan."

RESOURCES AND ENVIRONMENT

FACILITIES: Mark the facilities to be used at each performance site listed in Item 9, Face Page, and briefly indicate their capacities, pertinent capabilities, relative proximity, and extent of availability to the project. Use "Other" to describe the facilities at any other performance sites listed in Item 9 on the Face Page and at sites for field studies. Use continuation pages if necessary. Include an explanation of any consortium/contractual arrangements with other organizations.

☐ Laboratory: The Barany group currently numbers about a dozen full-time researchers, and two and a half 490 ft^2 laboratories (Kolthoff 476, 468, and 470; listed in order of length of occupancy; 4 desks per lab; common service corridor). These labs are down the hall from the faculty office. A 260 ft^2 laboratory (Kolthoff 463A, 1 desk) adjoins the office and is used for work by Dr. Barany and a laboratory technician and /or undergraduate research assistants. The group also occupies a 300 ft^2 instrumentation room (Kolthoff 472), and has a 125 ft^2 section of a shared instrumentation room (498A). All of this is in the Department of Chemistry on the Minneapolis campus of the University of Minnesota.

☐ Clinical:

☐ Animal:

☐ Computer: Three Macintosh personal computers for word processing and graphics

☐ Office: 140 ft^2 (Kolthoff 461)

☐ Other ( ):

MAJOR EQUIPMENT: List the most important equipment items already available for this project, noting the location and pertinent capabilities of each.

MilliGen/Biosearch 9050 and 9600 Peptide Synthesizers (acquired 1990); Beckman System 6300 High Performance Amino Acid Analyzer (acquired 1989); Beckman-Altex analytical gradient HPLC apparatus (acquired 1981) with variable wavelength UV detector and Hewlett-Packard recording integrator; Waters Delta-Prep 3000 HPLC (acquired 1988) apparatus with UV detector, integrator, and automatic sample injector; Beckman P/ACE 2100 capillary zone electrophoresis system (acquired 1991); MPLC set-up; UV-visible spectrophotometer; photolysis equipment; fraction collectors; Labconco lyophilizer; basic organic synthesis set-up; excellent hoods in all of the laboratories.

ADDITIONAL INFORMATION: Provide any other information describing the environment for the project. Identify support services such as consultant, secretarial, machine shop, and electronics shop, and the extent to which they will be available to the project.

Buildings of the Department of Chemistry contain major instrumental facilities for routine and high-field ^{1}H and ^{13}C-NMR, IR, mass spectrometry (including FABMS), and X-ray diffraction which are extensively used in this research program. Machine, electronics, and glassblowing shops are on-site, as is a research stockroom. The Department offers quarter-time secretarial support to the faculty, and Dr. Barany occasionally gets additional secretarial help paid for by research grants. An additional resource for this research is the Microchemical Facility of the University of Minnesota Institute of Human Genetics on the Minneapolis campus, which includes state-of-the-art equipment for amino acid analysis, peptide sequencing and synthesis, oligonucleotide synthesis, and other procedures.

A. SPECIFIC AIMS:

This Research Plan seeks to develop and optimize new tools that will be essential components to the integrated approach to cancer, genetic, and infectious disease detection described in the overall program project application. A novel polymerase chain reaction/ligase detection reaction (PCR/LDR) method for discriminating normal, carrier, and disease individuals has been described, and a high-sensitivity PCR/restriction endonuclease/LDR (PCR/RE/LDR) method for detection of rare cancer mutations is under development (Projects 1 and 2). These technologies, in their present form or as improved by applications of "convertide" nucleotide analogue bases (Project 3) and/or engineered thermostable ligase (Project 4), will be carried out in multiplex formats to simultaneously identify many mutations. LDR products, derived from one fluorescent primer and an adjacent primer with extra nucleotides or hexaethylene oxide "tails", are currently separated by size on an automated DNA sequencer, or by capillary electrophoresis. Use of different fluorescent groups allows a second dimension of mutation discrimination.

Herein, we propose new solid-phase approaches for simultaneous detection of multiplex LDR products. The general idea is that specific products will be captured on a spatially addressable array, so that the position of a signal identifies a mutation. Each LDR product will have a "zip code" tail, which will be captured selectively by a "complementary zip code" on the solid support. The supported (complementary) components can be modified DNA oligonucleotides or peptide nucleotide analogues (PNA), designed so that the resultant zip code hybrids have a significantly higher T_m than DNA/DNA hybrids. Unreacted primers may be washed away at high temperatures, allowing for detection of the LDR product. Multiple reuse of a universal "complementary zip code" array is envisaged to allow detection of a wide range of cancers and genetic diseases.

Implementation of these concepts, with the ultimate goal of achieving reliable and efficient materials and procedures that can be incorporated into easy-to-use, automated, low-cost diagnostic devices, will follow these aims:

(i) Development and evaluation of solid support materials compatible with chemical synthesis of DNA oligonucleotides and PNA oligomers, and compatible with subsequent hybridization reactions. Both commercially available and experimental materials will be screened. Surfaces, beads, or membranes will be functionalized, and extended as needed with hydrophilic spacers such as heterobifunctional polyethylene glycol (PEG) and/or carbohydrates. Chemistry for linking oligomers to the solid support, and/or solid-phase assembly of oligomers, will be developed.

(ii) Establishment of methodology for synthesis of spatially addressable arrays of DNA oligonucleotides and PNA oligomers. Appropriate masking technology will expose defined regions of the solid support for attachment of pre-formed oligomers, or for chain elongation to assemble the needed oligomers. In the latter mode, segment condensation will be used when possible in order to provide efficient convergent synthesis, and because chemical "failures" will become "invisible" during the subsequent hybridization.

(iii) Demonstration of scope and limitations of zip code concepts. As aims (i) and (ii) come to fruition, testing will be carried out (Core B). Design of primer and zip code structures will be facilitated by the informatics collaboration (Core A).

B. BACKGROUND AND SIGNIFICANCE

The cancer-detection technology of this program project application relies in good part on the capability to prepare by rapid, accurate chemical methods a multitude of oligonucleotide and related structures of defined sequence in the 20 to 50-base size range. The revolutionary solid-phase approach, introduced by Merrifield in the 1960's for peptides, points the way and provides ample experiences and precedents [1-6]. In solid-phase methodology that has been refined substantially and is readily automated, suitably protected amino acid building blocks are added in order (C to N) to a growing chain which is attached covalently through the C-terminus to a polymeric support. The principal commercially available

chemistries are referred to as "Boc"-and-"Fmoc", abbreviations for the names of the key *temporary* protecting groups. Often, a *linker* or *handle* is used to mediate the initial anchoring. Reactions are driven to completion by the use of excess reagents, which are removed by simple filtration and washing steps; the chemistry can also be carried out in columns in a continuous-flow mode. Upon completion of chain assembly, protecting groups are cleaved and the peptide is released into solution for further purification and characterization. An important aspect for successful results is the choice of the polymeric support. For many years, most work was carried out on 1% cross-linked microporous polystyrene resins (beaded), or on polyacrylamides (these latter could be embedded within an inorganic matrix, e.g., kieselguhr, or a rigid macroporous polystyrene, e.g., Polyhipe) [6-12]. Within the past few years, several additional materials with interesting physico-chemical properties have become available from several academic laboratories and commercial sources. These include membranes [13, 14], cotton and other appropriate carbohydrates [15-18], controlled-pore silica glass [19], and linear polystyrene grafted onto Kel-F [20]. A particularly interesting concept involves the use of polyethylene glycol-polystyrene supports (PEG-PS or Tentagel), which are compatible with both batchwise and continuous-flow reactors, and may facilitate difficult chemistries in peptide synthesis [21-23]. Other recent trends from the peptide field which are relevant to the goals of this proposal involve synthesis on polymeric surfaces, and the simultaneous preparation of multiple structures. In these procedures, due to Geysen, Frank, and research groups at Affymax, Arris, and Millipore, among others, relatively short peptides are built up on appropriately modified polyethylene pins, cellulose or polypropylene membranes, or glass surfaces, in a way that the bound final structures (purities in the 60 to 90% range) are tested directly by ELISA or other biological testing methods [15, 24-27]. The active structures are then deduced from their physical position, i.e., *spatial address* on the array. In an alternative combinatorial *library* approach developed at Selectide [28], millions of peptides are generated simultaneously by successive cycles of randomization/remixing of beaded supports. Intrinsic to the design of these experiments, each individual bead contains only a single peptide, so that those beads giving a positive interaction with a receptor can be picked out and subjected to analytical procedures that give the structure on the bead.

Solid-phase oligonucleotide synthesis has come to the fore in the past eight years with the development of reliable high-efficiency phosphoramidite [29] or H-phosphonate chemistry [30, 31] for linking protected nucleoside building blocks. Synthesis (3' to 5') is usually supported on controlled-pore glass, although other materials can also serve. The current automated methodology can routinely furnish oligonucleotides of length > 50 residues in overall purity > 90% directly upon release from the support. The methodology can also be adapted to incorporate unusual nucleotide bases, as well as modifications in the phosphodiester backbone (e.g., non-bridging thio or dithio substitution) and end-group labels (e.g., fluorescent dyes, biotin) [32]. As with peptides, relatively short oligonucleotides can be synthesized in spatially addressable arrays on glass surfaces [33-37]. Alternatively, a variety of procedures have been described for site-specific attachment of pre-synthesized oligonucleotide probes to nylon membranes or inside polyacrylamide gels [38-40]. Such arrays have been applied for DNA hybridization reactions, with applications to DNA sequencing or detection of biotinylated PCR-amplified products. These earlier studies provide useful precedents to some of the goals of this proposal.

Within the last few years, a group from Denmark [41-44] has introduced novel peptide nucleotide analogue (PNA) oligomers which mimic closely the spatial arrangement of the oligonucleotide backbone, but use nonchiral (2-aminoethyl)glycine units to replace the sugar phosphodiester (Figure 1). Additional innovations for PNA chemistry are under development at Millipore, in close consultation and collaboration with us [45]. These improvements include complete protection schemes for all the "bases" in concert with Boc, Fmoc, or alternative chemistries, optimized coupling (note: since racemization is not an issue, strong activation methods can be applied) and capping protocols, and efficient sequencing procedures. PNA and single-stranded DNA join to form anti-parallel heteroduplexes that exhibit Watson-Crick specificity and (particularly under low-salt conditions) tighter binding (higher T_m) than the corresponding double-stranded DNA [44]. As is discussed later, these properties dovetail extremely well with some of the requirements for the multiplex cancer detection protocol proposed in this program project application.

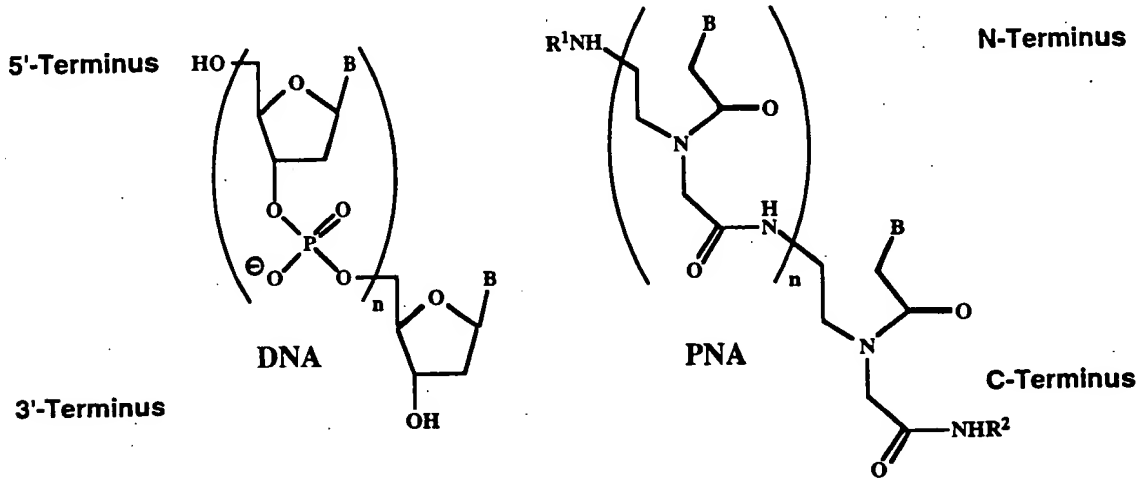


Fig. 1. Structural similarity of DNA and PNA.

C. PRELIMINARY STUDIES

The previous section of this proposal gave a brief overview of the current status of solid-phase methodologies, with an emphasis on general aspects which are relevant to the objectives of the research program. The following paragraphs provide brief descriptions of recent experimental advances from our laboratories which place us at the cutting edge of methodological developments.

We have invented several procedures for grafting both homo- and heterobifunctional polyethylene glycol (PEG) derivatives of defined molecular weight onto amino-functionalized polystyrene (PS); the resultant microporous beaded PEG-PS supports are now commercially available through Millipore and have numerous advantageous properties with respect to polystyrene [22, 23]. Our extensive experience in this field will be needed to introduce PEG as a "spacer" separating oligonucleotide or PNA molecules from functionalized surfaces. A useful start in this regard was reported recently by our collaborators Dr. Dereck Hudson and Dr. Ronald Cook [27] who in turn derivatized polyethylene plates, modified the resultant surfaces with PEG, and coupled carboxymethyl dextran to impart further hydrophilicity and serve as a starting point for peptide synthesis. (Please see letters of collaboration in overview section of program project).

PEG-PS has proven to be an ideal support for peptide library studies by the Selectide process [28]. PEG-PS is compatible with the organic reagents and solvents for efficient synthesis throughout the beads, and also has sufficient hydrophilic character to allow biological testing in aqueous milieus. Based on the realization that biological interactions occur only at the surface of beads, we have devised methodology for differentiation of "surface" and "interior" areas. PEG-PS is loaded with Boc-Trp-Gly, following which chymotryptic digestion "shaves" selectively only substrates at the surface which are accessible to the macromolecular enzyme. The exposed glycine is the starting point for orthogonal peptide synthesis using Fmoc chemistry, so that each bead is charged with a "screening" peptide at the surface, representing <1% of the total content but responsible for the entire spectrum of biological interaction. In concert, Boc chemistry establishes a sequenceable "coding" peptide confined to the interior, representing the vast majority of material on the bead but restricted from biological interaction [46]. This "shaving" concept has implications beyond the application to encoded combinatorial libraries just described; in the context of the present Research Plan, it can be used to ensure that oligopolymers synthesized on surfaces will be able to hybridize to oligonucleotides.

Our laboratories have also pioneered the development of novel linkers and handles for peptide synthesis [47, 48]; the extension to DNA and PNA is expected to be straightforward. Of particular interest, we have developed tris(alkoxy)benzyl amide (PAL) [49] and ester (HAL) [50] linkages, which upon cleavage with acid provide respectively C-terminal peptide amides, and *protected* peptide acids that can be

used as building blocks for so-called *segment condensation* approaches. We have noticed that the stabilized carbonium ion generated in acid from cleavage of PAL or HAL linkages can be intercepted by tryptophanyl-peptides. While this reaction is a nuisance for peptide synthesis and preventable (in part) by use of appropriate scavengers, we envisage herein a positive application to chemically "capture" oligo-Trp-end-labelled DNA and PNA molecules by HAL-modified surfaces (see Fig. 7 in Experimental Designs and Methods).

D. EXPERIMENTAL DESIGN AND METHODS

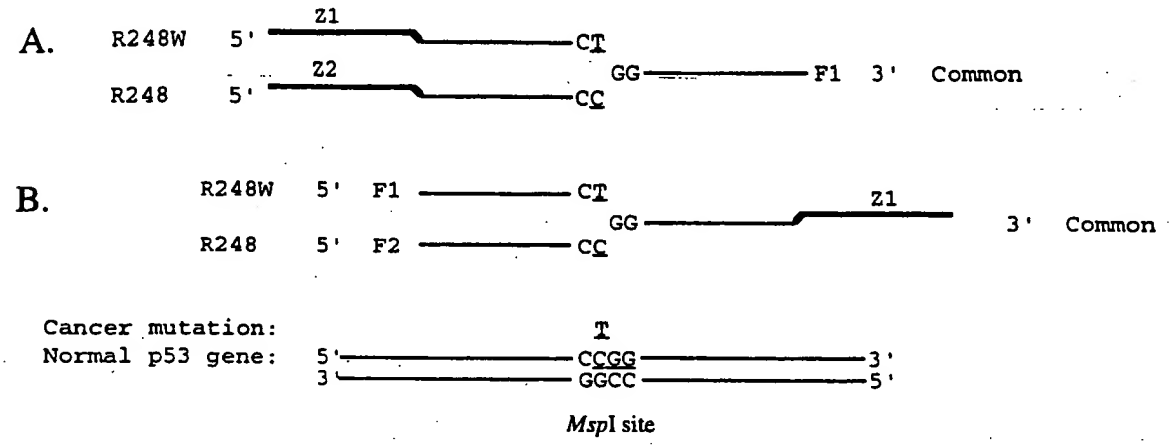
(i) Overview

(a) Array technology. This Reseach Plan describes a systematic approach to the design and synthesis of oligonucleotide or PNA arrays to achieve accurate detection and quantification of cancer mutations. Several groups have attempted to make oligonucleotide arrays with various degrees of success [33, 37- 40]. These approaches may be divided into three categories: (i) Synthesis of oligonucleotides by standard methods and their attachment one at a time in a spatial array [38-40] (ii) Photolithographic masking and photochemical deprotection on a silicon chip, to allow for synthesis of short oligonucleotides [37], and (iii) Physical masking to allow for synthesis of short oligonucleotides by addition of single bases at the unmasked areas [33, 36]. Although considerable progress has been made in constructing oligonucleotide arrays; some containing as many as 256 independent addresses, severe limitations have been noted in using these arrays for detecting specific DNA sequences by hybridizations. Arrays containing longer oligonucleotides can currently be synthesized only by attaching one address at a time, and thus are limited in potential size. (Current methods for attaching an oligonucleotide take about 1 hour, thus an array of 1,000 addresses would require over 40 days of around-the-clock work to prepare.) The "reverse dot blot" approach is capable of distinguishing single base differences in homozygous or heterozygous individuals, as well as the presence of a *ras* mutation diluted 20-fold by normal DNA [40]. However, hybridization methods require careful attention to temperature and salt conditions, and cannot achieve the high sensitivity of the cancer detection methods described in this proposal. Arrays containing large numbers of short oligonucleotides have performed significantly better on the computer than in practice. Syntheses on membranes or silicon chips are plagued by less than 100% efficiency, effectively limiting the size of these oligonucleotides to 8- to 10-mers. Imperfect hybridizations generate significant background signals, which severely hamper use of these arrays for DNA sequencing [36].

(b) Zip code concept. This proposal introduces a novel approach to oligonucleotide arrays which should obviate the above problems. One significant difference between our approach and literature array methods is that we use the array as a means to capture the *correctly generated* product. While others try to distinguish closely related sequences by subtle differences in melting temperatures during hybridization, we have already achieved the required exquisite specificity due to the discriminating actions of thermostable ligase in solution. Thus, our arrays can be designed to contain sequences which are *very different* from each other. Our array may be likened to 1,000 different antibodies that bind 1,000 different antigens with tight binding constants and no cross-reactivity. These arrays are completely universal, so that a single design may be used in detection of infectious and genetic diseases, or cancers. Best of all, the arrays will be highly stable and reusable.

The 1,000 different "antigens" are unique 24-mer "zip code" sequences linked covalently to the approximately 20- to 25-mer target-specific sequence of an LDR primer. A "zip code" sequence does not have any homology to either the target sequence or to other sequences on the genome. This zip code tail is then captured by its "antibody", a sequence complementary to the zip code on the addressable solid support array. The concept is shown in two possible formats for detection of the p53 R248 mutation (Fig. 2). At the top of the diagram shows two alternative formats for primer design to identify the presence of a germ line mutation in codon 248 of the p53 tumor suppression gene. The wild type sequence codes for arginine (R248) while the cancer mutation codes for tryptophan (R248W). The lower part of the diagram is a schematic for zip code capture. In the first format (A), the discriminating primers contain the allelic specific base (T for mutant and C for wild type) on the 3' end and 24-mer zip codes Z1 and Z2 on their 5' ends respectively. A common downstream primer contains a fluorescent group F1 at its 3' end. In the presence

presence of appropriate target DNA (wild type DNA is shown), the correct ligation products form. After hybridization of the zip code primers to their complementary sequences on the addressable array, unreacted fluorescent primers will be washed away. Mutant and wild type signal may be quantified using a FluorImager, and distinguished by their position on the array. In an alternative format (B), the discriminating oligonucleotides contain two different fluorescent groups F1 and F2, while the common oligonucleotide contains the zip code Z1. In this format, mutant and wild type signal are distinguished by the differences in fluorescence between F1 and F2 (see legend of Fig. 2 for more details).



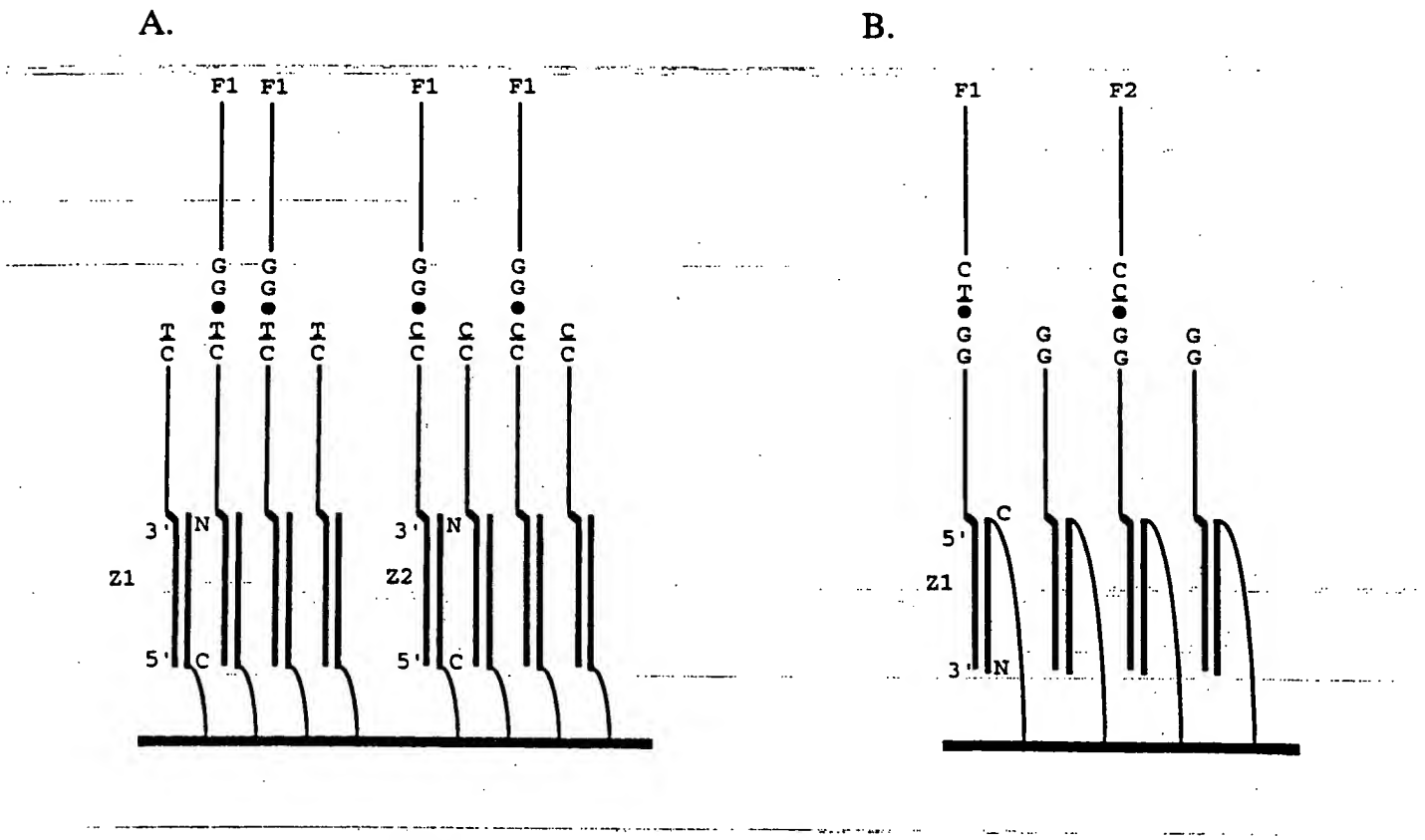


Fig. 2. Two alternative formats for zip code capture (see following page for legend).

Two alternative formats for zip code capture (legend for previous page. The top portion of the diagram shows two alternative formats for primer design to identify the presence of a germ line mutation in codon 248 of the p53 tumor suppressor gene. The wild type sequence codes for arginine (R248), while the cancer mutation codes for tryptophan (R248W). The bottom part of the diagram is a schematic diagram of zip code capture. The thick horizontal line depicts the membrane or solid surface containing the addressable array. The thin curved lines indicate a flexible linker arm. The thicker lines indicate a PNA sequence, attached to the solid surface in the C to N direction. For illustrative purposes, the PNA oligonucleotides are drawn vertically, making the linker arm in section B appear "stretched". Since the arm is flexible, the oligonucleotide will be able to hybridize 5' to C and 3' to N in each case, as dictated by base pair complementarity. A similar orientation of DNA/PNA hybridization would be allowed if the PNA were attached to the membrane at the N-terminus. Similar considerations apply when the complementary zip code on the support is a DNA oligonucleotide rather than PNA. (A) Two LDR primers are designed to discriminate wild type and mutant p53 by containing the discriminating base C or T at the 3' end. In the presence of the correct target DNA and *Tth* ligase, the discriminating primer is covalently attached to a common downstream oligonucleotide. The downstream oligonucleotide is fluorescently labeled. The discriminating oligonucleotides are distinguished by the presence of a unique "zip code" sequences, Z1 and Z2, at each of their 5' ends. A black dot indicates that target dependent ligation has taken place. After ligation, all zip code primers may be captured by their complementary "zip code" sequences at unique addresses on the array. Both ligated and unreacted primers are captured by the PNA array. Unreacted fluorescently labeled common primers and target DNA are then washed away at a high temperature (approximately 65°C to 80°C) and low salt. Mutant signal is distinguished by detection of fluorescent signal at the Z1 position, while wild type signal appears at the Z2 position. Heterozygosity is indicated by equal signals at both Z1 and Z2. The signals may be quantified using a Molecular Dynamics FluorImager. This format uses a unique address for each allele, and may be preferred for achieving very accurate detection of low levels of signal (30 to 100 attomoles of LDR product). (B) In this format, the discriminating oligonucleotides are distinguished by having different fluorescent groups, F1 and F2, on their 5' end. Either oligonucleotide may be ligated to a common downstream oligonucleotide containing a zip code sequence Z1 on its 3' end. In this format, both wild type and mutant LDR products are captured at the same address on the array, and are distinguished by their different fluorescence. This format allows for a more efficient use of the array and may be preferred when trying to detect hundreds of potential germline mutations.

(c) Design and synthesis of arrays. This Research Plan will explore variations of two general approaches for synthesizing arrays. In the first approach, we will prepare full-length 24-mer DNA oligonucleotides or PNA oligomers, which are subsequently linked covalently to a solid support or membrane. Alternatively, the deprotected DNA or PNA may remain linked to the bead, and the entire bead glued to a solid support. In the second approach, 36 specially designed PNA tetramers will be synthesized. These tetramers will be added to specific rows or columns on a solid support or membrane surface. The resulting "checkerboard" pattern will generate unique addressable arrays of PNA 24-mers.

We will initially explore glass and derivatized membrane supports to test their sensitivities and capacities as array surfaces. Pilot experiments will involve synthesis of five zip code PNA oligomers or oligonucleotides (sequences listed in Table 2, later). These oligomers will be covalently linked to the test surfaces. Fluorescently labeled complementary DNA zip code sequences will be synthesized in Core B, and used for testing arrays produced in this project. Note that for the initial studies, we do not require the longer conjugates that combine the LDR primer with the zip code (see Fig. 1).

What properties are desired in an array? The most important factor is good loading of oligonucleotide or PNA oligomer in a relatively small, but well-defined area. The current commercially available fluorescent imager can detect a signal as low as 2 attomoles per 50μ square pixel. Thus, a reasonable size address or "spot" on an array would be about 4 x 4 pixels, or 200μ square. The limit of detection for such an address would be about 32 attomoles per "spot", which is comparable to the 100 attomole detection limit using a DNA sequencing machine. The capacity of oligonucleotide which can be loaded per 200μ square will give an indication of the potential signal to noise ratio. A loading of 20 fmoles would give a signal to noise ratio of 625 to 1, while 200 fmoles would allow for a superb signal to noise ratio of 6,250 to 1. Loadings in excess of 200 fmoles will be unnecessary, since most LDR reactions use only 200 fmoles of each primer. The oligonucleotide or PNA oligomer should be on a flexible "linker arm" and on the "outside" or "surface" of the solid support for easier hybridizations. The support should be non-fluorescent, and should not interfere with hybridization nor give a high background signal due to nonspecific binding. In a mode where bead(s) are attached (typical size 50-200μ), neither beads nor "glue" should give a high background signal due to nonspecific binding or intrinsic fluorescence.

This proposal also introduces a novel approach for the design and synthesis of a universal PNA oligonucleotide array with 1,296 addresses. We envision each address to be about 200μ with an equal size

space in between addresses. Feasibility will be assessed with a 25 address array that is about 2 mm square; the full-sized array would be about 1.4 cm square. Preparation of such arrays (pilot and full-sized) will be carried out as a joint project with our industrial collaborators Dr. James Coull and his team at Millipore, and Dr. Ronald Cook who heads Siris Labs. (Please see letters of collaboration in overview section). The required preliminary synthesis in the academic laboratories will use a Biorad dot blot apparatus which contains individual microtiter wells sandwiched around a membrane. This allows for addition and filtration of chemicals in each well.

(ii) Design and optimization of zip codes and addresses

(a) General considerations. The principle of using zip codes has been explained earlier. The complementary zip codes (addresses) on the solid supports can be either DNA or PNA. *Both* will be tested. However, we expect that PNA-based capture of zip codes may have advantages over DNA-based capture because PNA/DNA duplexes are much stronger than DNA/DNA duplexes, by about 1°C/base-pair [44]. Thus, for a 24-mer DNA/DNA duplex with T_m = 72°C, the corresponding duplex with one PNA strand would have a "predicted" T_m = 96°C (the actual melting point might be slightly lower as the above "rule of thumb" is less accurate as melting points get over 80°C). Additionally, the melting difference between DNA/DNA and PNA/DNA becomes even more striking at low salt.

(b) Enhancement of the hybridization affinity of zip code/address duplexes. The melting temperature of DNA/DNA duplexes can be estimated as $[4n(G•C) + 2m(A•T)]$°C. If possible, we would like to narrow the T_m difference between zip code duplexes resulting from differences in G•C/A•T content, and in this way further optimize zip code capture. Froehler has shown that use of 5-propynyl-dU in place of thymine increases the T_m of DNA duplexes an average of 1.7°C per substitution [51]. We suggest that the same substitution in the zip code capture scheme would lower the T_m difference between zip code/address duplexes, and raise the T_m for all of the zip code/address duplexes. Phosphoramidite derivatives of 5-propynyl-dU (Fig. 3) will be prepared according to Froehler [51]. The 5-propynyluracil PNA monomer with Fmoc amino protection will be made (Fig. 4) following the published synthesis of PNA monomers [41, 42], replacing thymine with 5-iodouracil and using Pd(0) coupling of the alkylated 5-iodouracil and propyne. These monomers will be incorporated into synthetic DNA and PNA strands, respectively, and evaluated as described later.

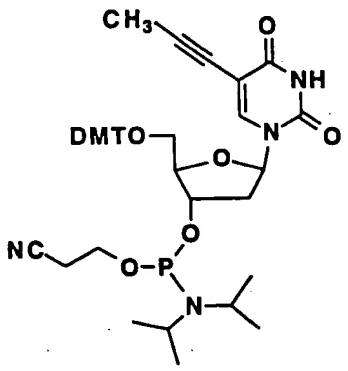
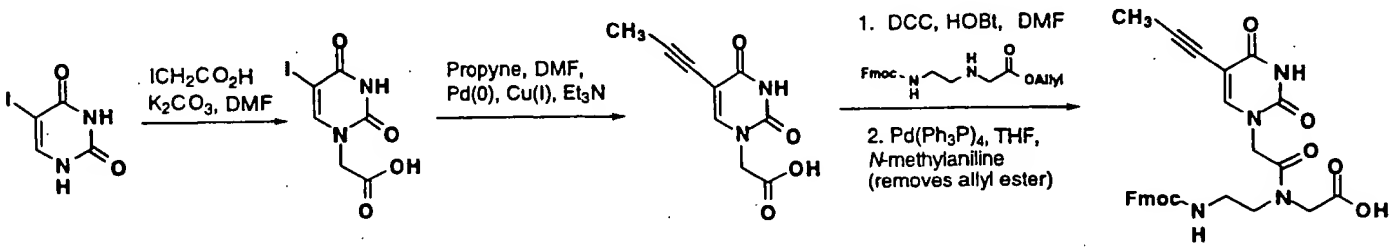


Fig. 3. Structure of nucleoside analogue 5-propynyl-dU.



The Boc-protected derivative could be made by a similar route.

Fig. 4. Synthesis of Fmoc-protected 5-propynyl-uridine PNA monomer.

(c) Zip code sequences designed from tetramer building blocks. Of the 256 (4^4) possible ways in which four bases can be arranged as tetramers, we have selected 36 that have unique sequences (Fig. 5). Each of the chosen tetramers differs from all the others by at least two bases, and no two dimers are complementary to each other. Furthermore, tetramers that would result in self-pairing or hairpin formation of the addresses have been eliminated (see legend to Fig. 5 for further details of the design process).

The final tetramers are listed in Table 1, and have been numbered arbitrarily from 1 to 36. Our premise is to use this unique set of tetramers as design modules for the required 24-mer zip code and 24-mer address sequences. The structures can be assembled by stepwise (one base at a time) or convergent (tetramer building blocks) synthetic strategies. Note that the numbering scheme for tetramers allows us to abbreviate each zip code as a string of six numbers (e.g., second column of Table 2, in following section).

(d) Initial zip code test sequences. The concept of zip code 24-mers designed from a unique set of 36 tetramers (Table 1) allows a huge number of possible structures, $36^6 = 2{,}176{,}782{,}336$. We have chosen five structures (Table 2) that have nearly equal G + C content as targets for the graded set of studies that are needed to establish the proposed methodology.

(e) Solution studies of zip code annealing. Our eventual goal is to exploit zip code hybridization to direct fluorescently labelled LDR products towards specific addresses on a solid support. However, first we wish to validate duplex formation in solution. Test sequences (Table 2) will be used. The T_m of each duplex will be measured by recording the A_{260} of the oligonucleotide solutions (~ 5μM concentration of each single strand) versus temperature.

Synthetic probes (normal and complementary directions) for the aforementioned studies will be prepared as either DNA or PNA, with either all thymine or all 5-propynyl-uracil. Where syntheses are straightforward, they will be performed by Core B, but where methodology is still under development, syntheses will be performed in the laboratories of program project chemists or industrial collaborators. These syntheses will generate for each sequence a total of eight oligomers, which can be combined in 16 ways that form duplexes.

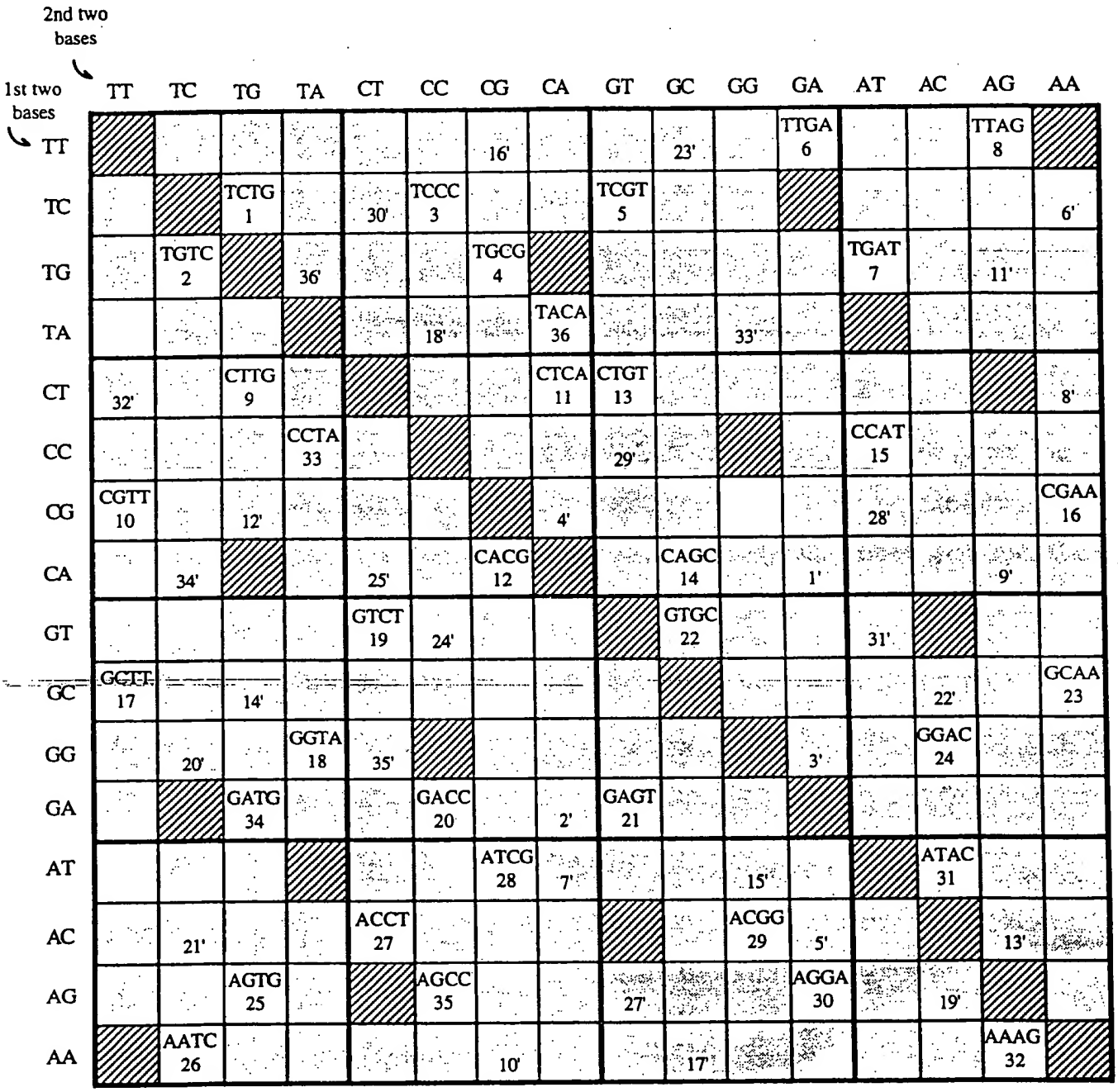


Fig. 5. Design of 36 tetramers which differ from each other by at least 2 bases. Checkerboard pattern shows all 256 possible tetramers. A given square represents the first two bases on the left followed by the two bases on the top of the checkerboard. Each tetramer must differ from each other by at least two bases, and should be non-complementary. The tetramers are shown in the white boxes, while their complements are listed as (number)'. Thus, the complementary sequences GACC (20) and GGTC (20') are mutually exclusive in this scheme. In addition, tetramers must be non-palindromic, e.g., TCGA (darker diagonal line boxes), and non-repetitive, e.g., CACA (darker diagonal line boxes from upper left to lower right). All other sequences which differ from the 36 tetramers by only 1 base are shaded in light gray. Four potential tetramers were not chosen as they are either all A•T or G•C bases. In addition, thymine can be replaced by 5-propynyl uridine when used within DNA or PNA address sequences as well as in the DNA zip code sequences. This would increase the T_m of an A•T base pair by ~1.7°C. Thus, T_m values of individual tetramers should be approximately 15.1°C to 15.7°C. T_m values for the full length 24-mers should be 95°C or higher.

Table 1. List of tetramer PNA sequences and complementary DNA sequences, which differ from each other by at least 2 bases.

Number	Sequence (N-C)	Complement (5'-3')	G + C
1.	TCTG	CAGA	2
2.	TGTC	GACA	2
3.	TCCC	GGGA	3
4.	TGCG	CGCA	3
5.	TCGT	ACGA	2
6.	TTGA	TCAA	1
7.	TGAT	ATCA	1
8.	TTAG	CTAA	1
9.	CTTG	CAAG	2
10.	CGTT	AACG	2
11.	CTCA	TGAG	2
12.	CACG	CGTG	3
13.	CTGT	ACAG	2
14.	CAGC	GCTG	3
15.	CCAT	ATGG	2
16.	CGAA	TTCG	2
17.	GCTT	AAGC	2
18.	GGTA	TACC	2
19.	GTCT	AGAC	2
20.	GACC	GGTC	3
21.	GAGT	ACTC	2
22.	GTGC	GCAC	3
23.	GCAA	TTGC	2
24.	GGAC	GTCC	3
25.	AGTG	CACT	2
26.	AATC	GATT	1
27.	ACCT	AGGT	2
28.	ATCG	CGAT	2
29.	ACGG	CCGT	3
30.	AGGA	TCCT	2
31.	ATAC	GTAT	1
32.	AAAG	CTTT	1
33.	CCTA	TAGG	2
34.	GATG	CATC	2
35.	AGCC	GGCT	3
36.	TACA	TGTA	1

(iii) Solid support materials for array technology

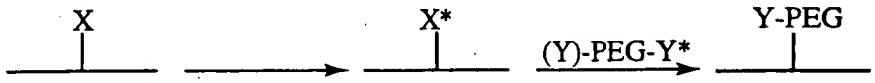
Earlier sections of this proposal have reviewed the ideal requirements for array support materials, in the context of options from the literature and our own extensive experiences. The solid supports must be charged with DNA oligonucleotides or PNA oligomers; this is achieved either by attachment of pre-synthesized probes, or by direct assembly and side-chain deprotection (without release of the oligomer) onto the support. Further, the support environment needs to be such as to allow efficient hybridization. Towards this end, three factors may be identified: (i) sufficient hydrophilic character of support material (e.g., PEG or carbohydrate moieties); (ii) flexible linker arms (e.g., hexaethylene oxide or longer PEG chains) separating the probe from the support backbone; (iii) "shaving" procedures which allow probe immobilization or probe synthesis to occur only in the most accessible "surface" areas of the support. It should be kept in mind that numerous ostensibly "flat surfaces" are quite thick at the molecular level. Lastly, it is important that the support material not provide significant background signal due to non-specific binding or intrinsic fluorescence.

Table 2. List of initial PNA zip code and complementary DNA oligonucleotides.

Polymer	Zip code	Sequence			G+C
		16 3 34 2 9 1			
PNA	16-3-34-2-9-1	NH_2- CGAA-TCCC-GATG-TGTC-CTTG-TCTG	-COOH		13
DNA	1-9-2-34-3-16(c)	5'- CAGA-CAAG-GACA-CATC-GGGA-TTCG	-3'		13
		7 3 11 2 18 1			
PNA	7-3-11-2-18-1	NH_2- TGAT-TCCC-CTCA-TGTC-GGTA-TCTG	-COOH		12
DNA	1-18-2-11-3-7(c)	5'- CAGA-TACC-GACA-TGAG-GGGA-ATCA	-3'		12
		20 3 14 2 7 1			
PNA	20-3-14-2-7-1	NH_2- GACC-TCCC-CAGC-TGTC-TGAT-TCTG	-COOH		14
DNA	1-7-2-14-3-20(c)	5'- CAGA-ATCA-GACA-GCTG-GGGA-GGTC	-3'		14
		29 3 23 2 12 1			
PNA	29-3-23-2-12-1	NH_2- ACGG-TCCC-GCAA-TGTC-CACG-TCTG	-COOH		15
DNA	1-12-2-23-3-29(c)	5'- CAGA-CGTG-GACA-TTGC-GGGA-CCGT	-3'		15
		13 35 27 33 2 7			
PNA	13-35-27-33-2-7	NH_2- CTGT-AGCC-ACCT-CCTA-TGTC-TGAT	-COOH		12
DNA	7-2-33-27-35-13(c)	5'- ATCA-GACA-TAGG-AGGT-GGCT-ACAG	-3'		12

A variety of materials, which include suitably modified glass, plastic, or cellulose surfaces, PEG-PS beads, or a variety of membranes, will be examined in the context of the needs summarized above. These materials will be obtained from commercial sources or from our industrial collaborators (Dr. James Coull at Millipore, Dr. Ronald Cook at Siris), or else will be prepared in our laboratories by following literature precedents. Depending on the material, surface functional groups (i.e., hydroxyl, carboxyl, amino) may be present from the outset (perhaps as part of the coating polymer), or will require a separate procedure (e.g., plasma amination, chromic acid oxidation, treatment with a side-chain functionalized alkyltrichlorosilane) for introduction of the functional group. Hydroxyl groups become incorporated into stable carbamate (urethane) linkages by several methods. Amino functions can be acylated directly, whereas carboxyl groups are activated, e.g., with N,N'-carbonyldiimidazole or water-soluble carbodiimides, and reacted with an amino-functionalized compound (Fig. 6). Unreacted amino groups will be blocked by acetylation or succinylation, to ensure a neutral or negatively charged environment that "repels" excess unhybridized DNA. Loading levels will be determined by standard analytical methods [47].

Often, it will be desirable to introduce a PEG spacer with complementary functionalization, prior to attachment of the starting linker for DNA or PNA synthesis. The methodology to do so is in hand [22, 23, 52] and will be pursued alongside with control experiments on the same materials lacking PEG. Similarly, dextran layers can be introduced as needed by precedented chemistries [27, 53]. Finally, enzymatic "shaving" is carried out readily by our recently developed procedure using chymotrypsin to cleave a short substrate that is distributed uniformly throughout a bead or on a derivatized surface. In our studies on peptide/receptor (antibody or binding protein) interactions, we have shown that shaving protocols expose a relatively small portion (approximately 1 to 5%) of the total functional groups, yet they reach all receptor-accessible sites [46]. We plan to establish whether the same site selectivity can be achieved for hybridization reactions, and compare the results to controls run on "unshaven" materials.

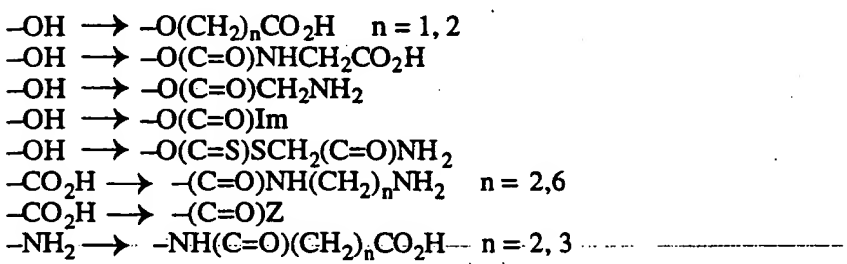


X,Y = –OH, $–CO_2H$, $–NH_2$
(Y) = W–NH–

X*, Y* = $–O(C=O)Z$, $–O(C=S)Z$, $–CO_2H$, $–(C=O)Z$, $–NH_2$, $–N=C=O$

W = protecting group, e.g. Boc, Fmoc
Z = activating group, e.g. imidazole (Im), *p*-nitrophenol (OPnp), hydroxysuccinimide (OSu), pentafluorophenol (OPfp)
PEG = oligo or poly(ethylene glycol), backbone $(CH_2CH_2O)_n$ n = 6 to 200 (can also be grown by anionic polymerization with )
WSC = water soluble carbodiimide

Functional group transformations/activation (as needed), X → X*, Y → Y*



Covalent linkage, X* + Y*

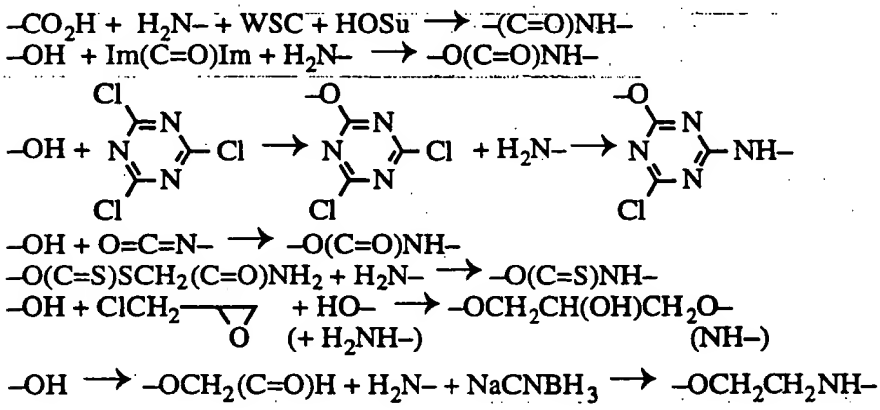


Fig. 6. Chemical reactions for covalent modifications, grafting, and oligomer attachments to solid supports. The solid supports can be beads, membranes, or surfaces, with a starting functional group X. Functional group transformations can be carried out in a variety of ways (as needed) to provide group X* which represents one partner in the covalent linkage with group Y*. The Figure shows specifically the grafting of PEG, but the same repertoire of reactions can be used (however needed) to attach carbohydrates (with hydroxyl), linkers (with carboxyl), and/or DNA oligonucleotides and PNA oligomers that have been extended by suitable functional groups (amino or carboxyl). In some cases, group X* or Y* is pre-activated (isolable species from a separate reaction); alternatively, activation occurs in situ. Referring to PEG as drawn in the Figure, Y and Y* can be the same (homobifunctional) or different (heterobifunctional); in the latter case, (Y) can be protected for further control of the chemistry.

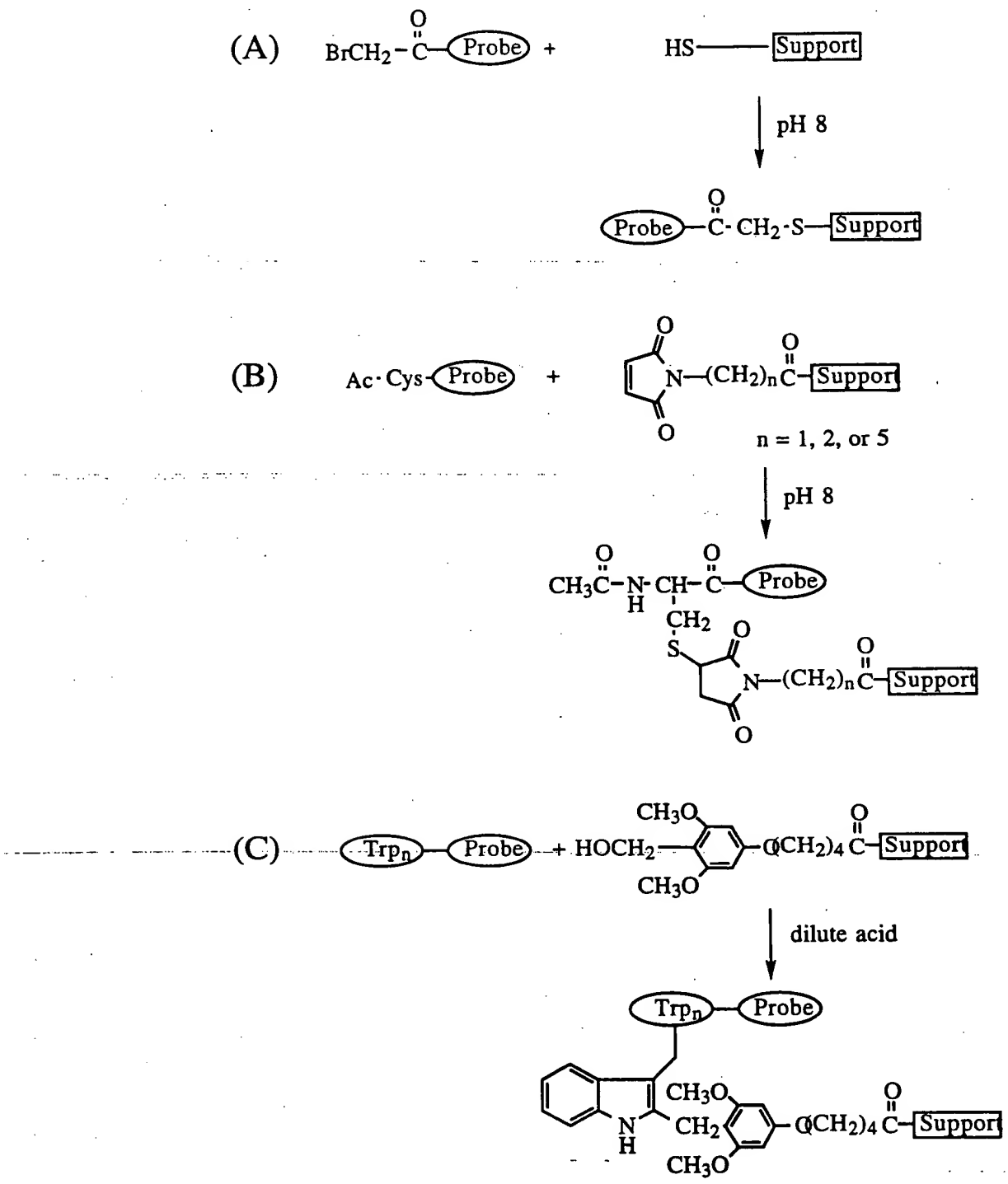


Fig. 7. Proposed chemistries for covalent attachment of DNA or PNA probes to solid supports. Chemically synthesized probes can be extended, on either end (shown here on the N-terminal of PNA or the 5' end of DNA). Further variations of the proposed chemistries are readily envisaged. **(A)** An amino group on the probe is modified by bromoacetic anhydride; the bromoacetyl function is captured by a thiol group on the support. **(B)** An N-acetyl, S-tritylcysteine residue coupled to the end of the probe provides, after cleavage and deprotection, a free thiol which is captured by a maleimido group on the support. **(C)** The probe contains an oligo-tryptophanyl tail (n = 1 to 3), which is captured after treatment of a HAL-modified support with dilute acid.

(iv) Immobilization of individually synthesized DNA oligonucleotides or PNA oligomers onto solid supports

(a) Synthesis of 24-mer DNA oligonucleotides and 24-mer PNA oligomers. The five sequences listed in Table 2 are designed for testing the hybridization properties of zip code arrays. They will be synthesized by Core B as DNA oligonucleotides using standard phosphoramidite chemistries [29], and incorporating an "aminolink" group at the 5'-terminus. In addition, the same sequences will be synthesized as PNA oligomers by stepwise Boc or Fmoc solid-phase chemistry, or by a segment condensation approach using suitably protected PNA tetramers. The PNA will have one endgroup blocked (e.g., acetyl on N-terminus, or amide on C-terminus), and the other terminus extended with ε-aminocaproic acid to provide a free aliphatic amino or carboxyl site for ultimate linking to the solid support. The complementary zip code oligonucleotide sequences will be prepared with a fluorescent "Fam" group at the 5'-end, by Core B.

Synthetic DNA oligonucleotides or PNA oligomers will be released from the resin supports, concurrent with removal of side-chain protecting groups. These modified oligomers will be purified to homogeneity by well-precedented polyacrylamide gel electrophoresis (PAGE) or high performance liquid chromatography (HPLC: reversed-phase or anion-exchange) procedures. Immobilization to solid supports will follow, as described below.

(b) Covalent attachment of DNA oligonucleotides or PNA oligomers to solid supports. The purified oligomers all contain a free aliphatic amino group at the terminus, which allows attachment to a derivatized membrane according to Zhang [40]. Other attachment chemistries based on amino group chemistry will also be pursued, building on a wealth of precedents for connecting functionalized polymers and proteins to each other and to solid matrices [54] (Fig. 6). These procedure can be carried out in series with several probes, resulting in site-specific attachments. Once the complementary zip code probes have been immobilized, oligonucleotide hybridizations using fluorescently labeled zip codes will be carried out to evaluate both capacity and signal to noise ratio, as described in Core B.

Encouraging results in the experiments outlined above would provide impetus to the exploration of additional immobilization ("capture") chemistries, which need to be rapid, specific, and non-destructive to the combination of functional groups found in DNA oligonucleotides and PNA oligomers. Our strategy involves incorporation, through synthesis, of alternative functional groups at either end of the probe, together with modification of the support by a suitable complementary functional group. More specifically, we can take advantage of the facile S-alkylation or Michael addition of thiol groups, or of the reaction in dilute acid of indole moieties with tris(alkoxy)benzyl carbocations (Fig. 7).

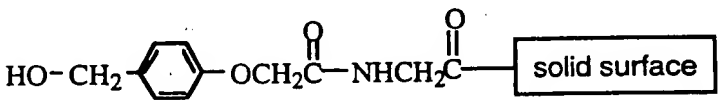
(v) Synthesis of oligonucleotides or PNA oligomers on solid supports and creation of arrays on solid surfaces

(a) Synthesis of 24-mer DNA oligonucleotides and 24-mer PNA oligomers on "shaved" beads. A second approach to constructing the arrays required for zip code capture starts with the assumption that suitable probes can be assembled and side-chain deprotected with covalent retention on beads used for solid-phase synthesis; these beads are then delivered to discrete addresses on a solid surface. General considerations have been outlined earlier; we consider this mode to represent a particularly pertinent case where application of our "shaving" concept may be critical to success. As before, the five sequences listed in Table 2 will be synthesized, either as DNA oligonucleotides (standard phosphoramidite chemistry) or as PNA oligomers (stepwise Boc or Fmoc chemistry).

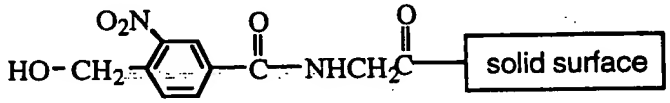
PEG-PS beads of 100μ diameter have a normal capacity of approximately 30 pmol, meaning that a shaved bead is predicted to hold about 0.1 to 0.5 pmol of final product. This level of material is well within the requirements of the subsequent hybridization studies. Chymotryptic "shaving" of a Boc-Trp-Gly-PEG-PS sequence generates a free α-amino group from "surface" glycine residues. For PNA synthesis, the C-terminal monomer is coupled to form a non-cleavable peptide bond; for DNA synthesis, an N-acetyl-serine spacer is introduced so that phosphoramidite synthesis begins off the free hydroxyl side-chain. For DNA chains upon completion of chain assembly, removal of the usual base-labile side-chain and phosphate protecting groups with aqueous ammonia yields the free probe oligomer linked covalently to the outside

areas of PEG-PS. For PNA oligomers, benzyloxycarbonyl-type protecting groups will be removed with strong acid, e.g., trifluoromethanesulfonic acid (the strategy may change if/when milder protection schemes under development at Millipore or by us are established).

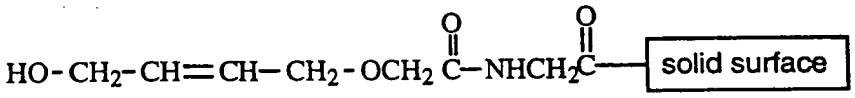
In order to better document the chemistry of chain assembly on "shaved" as well as control beads, the synthesis outlined above can be modified by introduction of base-stable, orthogonally cleavable linkers to separate the free glycine from the 3' or C-terminal residue. Suitable choices (Fig. 8) include acid-labile *p*-alkoxybenzyl (PAB), photolabile *o*-nitrobenzyl (ONb), or Pd(0)-cleavable allyl (Al) [47, 48]. The oligonucleotide or PNA products can be released from the support in a discrete step, and evaluated by standard analytical criteria as well as solution hybridization with the complementary sequences.



PAB



ONb



Al

Fig 8. Handles for attachment of oligopolymers to "shaved" beads. Handles are coupled through their carboxyl groups to "shaved" beads. The free hydroxyl on the left side of each structure can be esterified with the C-terminal PNA monomer or phosphitylated with a nucleoside phosphoramidite.

(b) Attachment of beads to solid surfaces. DNA or PNA synthesis on PEG-PS using automated instrumentation is generally carried out with several hundred thousand to several million beads, setting the stage for economical mass production of arrays. Each 200μ square pixel will contain several 100μ beads, and different spatial addresses will contain different beads. Several of our industrial collaborators are actively pursuing concepts for attaching beads to solid surfaces in a rapid and automated manner. (Please see letters of collaboration from Dr. James Coull of Millipore and Dr. Ronald Cook of Siris Labs). Briefly, these concepts include precise positioning of the bead over a gridded surface and melting one surface to attach them, using grooves or dimples in the surface to help position beads (using vacuum suction to guide a bead into a particular position), and/or projecting the beads onto a surface containing glue. For academic purposes, we will place beads onto a thin layer of some bonding material, such as epoxy. All bonding materials will need to be tested for resistance to high temperatures/high salt conditions, and to confirm the absence of non-specific binding to DNA oligonucleotides. Oligonucleotide hybridizations using fluorescently labeled complementary zip codes will be used to evaluate both capacity and signal to noise ratio as described in Core B.

(c) Variations. As necessary, the evolving literature methods for simultaneous syntheses of peptides or oligonucleotides at defined positions will be adapted in concert with the chemistry (linkers, protection strategies) and other concepts (shaving) presented above for beads. For example, it will be of interest to learn whether shaving of membranes, PEG-modified polyethylene surfaces, or pins helps with synthesis and/or hybridization efficiency. The various literature protocols for multiple synthesis are quite labor-intensive, but they may need to be pursued should we be unable to devise successful way to glue beads that contain DNA or PNA probes to surfaces, and to apply such materials for hybridization.

(vi) Direct synthesis of PNA arrays by masking/segment condensation on solid supports

(a) Perspective. In principle, arrays can be constructed most effectively by use of highly accurate masking and unmasking technology with per cycle yields of close to 100%. Unfortunately, current chemistries proceed in at best 97% yield per step, with a possible further drop-off as chain length increases. These relatively low efficiencies allow for construction of modest arrays in the octamer to decamer size range (even so, with substantial synthetic error rates which translate to hybridization at false addresses), but preclude construction of 24-mers needed as complementary zip codes in the cancer detection scheme of this program project proposal.

In the following, we propose a novel way to circumvent the aforementioned problems. The design of zip code arrays (in which individual arrays have substantial differences to minimize any chances of cross-reactivity; see earlier Fig. 5 and Table 1, and accompanying discussion) has been *integrated* with the synthetic strategy. Rather than carrying out stepwise synthesis to introduce bases one at a time, we use protected PNA tetramers as building blocks. (As shown below, these are easy to prepare; the corresponding protected oligonucleotide intermediates would require additional protection of the internucleotide phosphate linkages.) Construction of the 24-mer at any given address requires only six synthetic steps, with a likely improvement in overall yield by comparison to stepwise synthesis. Moreover, since failure sequences at each address are shorter and lacking at least four bases, there is no risk that these will interfere with correct hybridization or lead to incorrect hybridizations. This insight also means that "capping" steps will not be necessary.

Masking technology will allow several addresses to be built up simultaneously, as is explained below. As direct consequences of the manufacturing process for the arrays, several further advantages are noted. Each 24-mer address differs from its nearest 24-mer neighbor by three tetramers, or at least 6 bases. At low salt, each base mismatch in PNA/DNA hybrids decreases the melting temperature by 8°C. Thus, the T_m for the correct PNA/DNA hybridization is at least 48°C higher than any incorrect hybridization. Also, neighboring 24-mers are separated by 12-mers, which do not hybridize with anything and represent "dead" zones in the cancer detection profile. Finally, by choosing PNA addresses, we create rugged, reusable arrays.

The remaining description indicates methods for preparation of 36 unique PNA tetramers, and shows the mechanical/chemical strategy to prepare the arrays. Pilot experiments will result in the creation of a 5x5 array with 25 addresses of PNA 24-mers. Ultimately, all 36 tetramers can be incorporated to generate full-size arrays of 1,296 addresses.

(b) Synthesis of protected PNA tetramer building blocks. For each of the 36 unique sequences that have been designed (Table 1), we require the intermediates with appropriate protection on the α-amino group, on the side-chains, and with a free α-carboxyl group. This can be done readily by Fmoc chemistry on PAB or HAL resins, according to standard protocols published from our laboratory [50, 55]. Following cleavage in dilute acid, the protected intermediates will be purified by chromatography. Alternatively, we can use Boc chemistry with ONb (photolabile) or allyl (cleaved by Pd(0)) resins [47, 48, 56].

(c) Construction of PNA arrays. As stated already, only the pilot study with a 5x5 array is described. Considerations concerning the solid support are the same as described earlier: starting surfaces will contain free amino groups ("shaved" if necessary), a non-cleavable amide linkage will connect the C-terminus of PNA to the support, and orthogonal side-chain deprotection must be carried out upon completion of segment condensation assembly in a way that PNA chains are retained at their addresses. A simple masking device has been designed that contains 200μ spaces and 200μ barriers, to allow each of 5 tetramers to couple to the solid support in distinct rows (Fig. 7). After addition of the first set of tetramers, the masking device is rotated 90°, and a second set of 5 tetramers are added. This can be compared to putting icing on a cake as rows, followed by icing as columns. The intersections between the rows and columns will contain more icing, likewise, each intersection will contain an octamer of unique sequence. Repeating this procedure for a total of 6 cycles generates 25 squares containing unique 24-mers, and the remaining squares containing common 12-mers (Fig. 10).

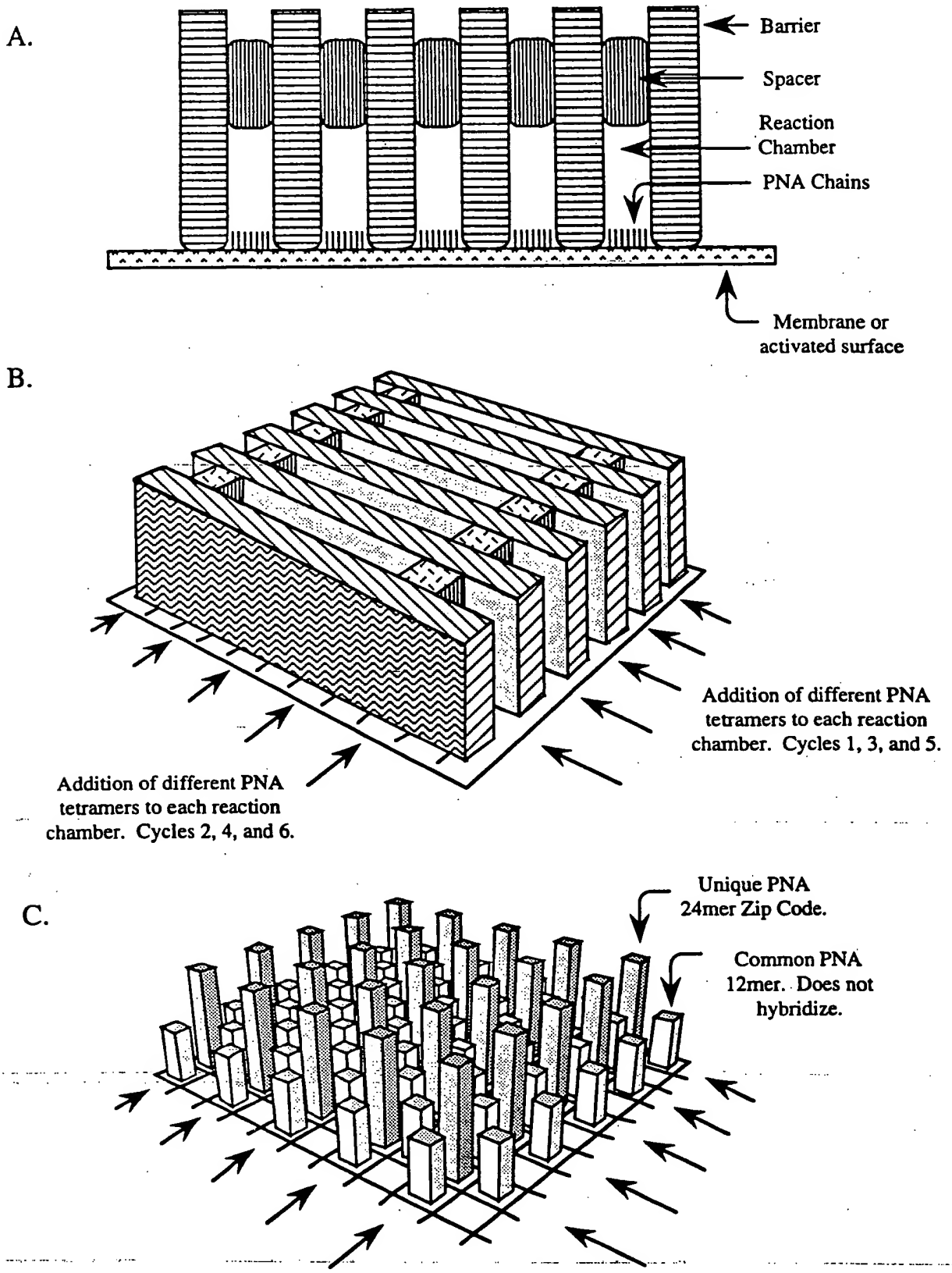


Fig. 9. Process for manufacturing an array (see following two pages for Fig. 10 and legends).

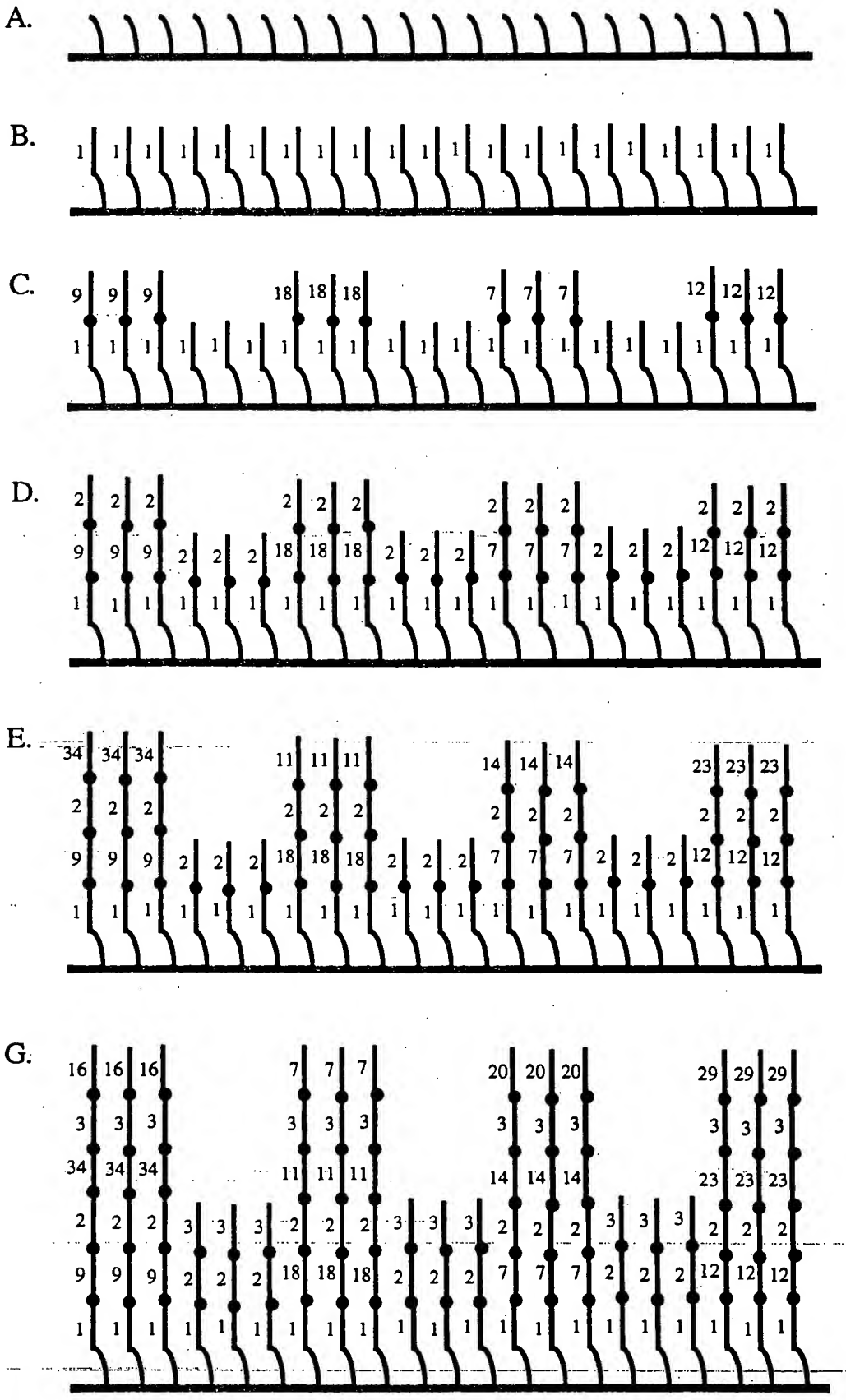


Fig. 10. Schematic cross-sectional view of synthesis of addressable array (see next page for legend).

Fig. 9. Process for manufacturing an array (legend). (A). Side view of reaction chambers. (B) Three-dimensional view of reaction chambers. Each wall and spacer is 100μ thick. These spacers form chambers of width 100μ. The multi-chamber device is pressed onto the membrane or activated solid surface, forming tight seals. The barriers may be coated with rubber or another material to avoid cross contamination from one chamber to the next. One must also make sure the membrane or solid support surface is properly wetted by the solvents. The membrane can be in a vertical position with a plate on each side to clamp the multi-chamber device to the membrane. Solvents are introduced at the bottom, rise up the chamber, and are removed from the top, much like an ABI four channel DNA synthesizer, except now there are 36 chambers and 36 different tetramer bottles. One proceeds by activating the surface, deprotecting, and adding a tetramer. The chamber is unclamped, the membrane is rotated 90°, and reclamped. A second round of tetramers are added. (C) Schematic bird's eye view of PNA oligomer array after completion of all 6 rounds of synthesis. Each tower represents 100 fmole of oligomers. Taller towers represent full size 24-mers which result from 6 rounds of synthesis in alternating directions. Each 24-mer tower represents a unique PNA sequence. Smaller towers represent half-size 12-mers which result from 3 rounds of synthesis in the same direction. All smaller towers in the same row as the arrows are of identical sequence. For clarity, the towers have been drawn as individual units, even though in the "real" synthesis they will be the same dimension as the grid squares and thus appear fused to each other. A "side view" of these individual towers is shown in Fig. 10.

Fig. 10. Schematic cross-sectional view of synthesis of addressable array (legend). (A) Attachment of flexible spacer (linker) to surface of array. (B) Synthesis of the first rows of PNA tetramers. Only the first row, containing tetramer 1, is visible. The multi-chamber device is placed so that additional rows, each containing a different tetramer, are behind the first row. (C) Synthesis of the first columns of PNA tetramers. The multi-chamber device has been rotated 90°. Tetramers 9, 18, 7, and 12 were added in adjacent chambers. (D) Second round synthesis of the PNA rows. The first row contains tetramer 2. (E) Second round of synthesis of PNA columns. Tetramers 34, 11, 14, and 23 were added in adjacent chambers during the second round. (F) (Not shown) Third round synthesis of PNA rows. The first row contains tetramer 3. (G) Structure of array after third round synthesis of columns, adding tetramers 16, 7, 20, 29. Note that all 24-mer PNA oligomers within a given row or column are unique, hence achieving the desired addressable array. Since each 24-mer differs from its neighbor by three tetramers, and tetramer differs from each other by at least 2 bases, then each 24-mer differs from the next by at least 6 bases. Each mismatch significantly lowers T_m, and the presence of 6 mismatches in just 24 bases would make cross hybridization unlikely even at 35°C. Note that the smaller 12-mer sequences are identical with one another, but are not at all common with the 24-mer sequences. Even though the particular 12-mer sequence may be found within a 24-mer elsewhere on the grid, for example 17-1-2-3-28-5, a zip code will not hybridize to the 12-mer at temperatures above 50°C.

Our design for a masking device is essentially the same as the masking technique developed by Maskos and Southern [33, 36]. This device will facilitate the desired array synthesis, and allow us to move on to testing zip code hybridization with Core B. The masking device will be designed and prepared by our industrial collaborators Dr. Ronald Cook of Siris Labs, and Dr. James Coull of Millipore. Simultaneously, we will also create, by hand, test arrays on membranes with aid of the Biorad dot blot apparatus containing individual microtiter wells.

E. PROGRAM ASPECTS

We are developing solid-phase methodology which will allow multiplex detection o: oligonucleotide ligation products that are indicative of cancer mutations. Specific aims of this project (Project 5) are: (i) Development and evaluation of solid support materials compatible with chemical synthesis of DNA oligonucleotides and PNA oligomers, and compatible with subsequent hybridization reactions. (ii) Establishment of methodology for synthesis of spatially addressable arrays of DNA oligonucleotides and PNA oligomers. (iii) Demonstration of scope and limitations of zip code concepts.

The zip code approach, including the key needs for its experimental implementation, arose throug extensive discussions between F. Barany, R.P. Hammer, and G. Barany. The studies described in Project 5 that are directed towards solid support development and evaluation, with respect to either immobilization of pre-synthesized oligomers or their direct synthesis, will interface closely with efforts of Core B to prepare the needed DNA and PNA primer, zip code, and complementary zip code sequences, and to carry out hybridization assays. A collaboration with Project 3 will provide 5-propynyl-uridine monomers for incorporation into either DNA or PNA that may have improved thermodynamic parameters in hybridization. Progress in Project 5 on the preparation and application of spatially addressable arrays to detection of LDR products will have an immediate impact on the cancer work described by Project 1 and 2, since it will then

be possible to test clinical samples at a significantly enhanced throughput. We also anticipate considerable interactions between Project 5 and Core A for computer-aided design of zip code sequences.

Addressable array capture will eventually be the preferred method of identifying mutations. In our initial examination of the p53 gene in colon, lung, and breast tumor samples we will only look for nine different mutations, V157, R175, H179, C242, G245, R248, R249, R273 and R282 (See Projects 1 and 2). By synthesizing LDR primers with tails of varying lengths we can easily distinguish between these mutations using gel or capillary electrophoresis. However, increasing the number of assayable mutations eventually makes electrophoretic detection less feasible. This is due to two reasons. First, mutation-specific LDR primers should differ in length by two bases for their products to be distinguished by electrophoresis. For a large number of mutations to be assayed together, very long primers would have to be synthesized. Second, only one mutant signal would be expected for most reactions. Since mutant signals will differ from each other by only two bases, minor defects in a gel lane could cause a misreading of the LDR product length and incorrect identification of the mutation. Both of these problems are currently overcome in our laboratory by using more than one color fluorescent label and internal standards within the same lane. However, once the primers become very long (75-100 bases), failure sequences (n-1, n-2) become increasingly harder to separate by HPLC or gel purification. In contrast, by synthesizing a unique 24 base zip code sequence to each LDR primer, the product can be captured by its complementary zip code at a discrete "address" on a two-dimensional array. Failure sequences do not present a problem for either the zip code sequence or its complementary address. A fluorecent signal at a specific address, as opposed to a specific size, thus indicates the presence of a specific cancer mutation. (See Core B.)

A reusable addressable array with high capacity and excellent signal to noise specificity would be of benefit to several of our collaborators who need to detect large number of mutations. This will aid in the detection and identification of: hundreds of microorganisms by identifying 16s polymorphisms (Dr. Carl Batt), dozens of ß-lactamase mutations responsible for third generation ß-lactam resistance (Dr. Patrice Courvalin), epidemiological studies based on HIV polymorphisms (Dr. Olen Kew), dozens of polymorphisms in the E6 and E7 genes of high risk HPV strains (Dr. Saul Silverstein), multiple germline mutations in single gene disorders (Dr. Eric Hoffman, Dr. Perry White, and Dr. Emily Winn-Deen), and multiple somatic mutations in tumor suppressor genes and oncogenes (Dr. John Kovach, Dr. Michael Osborne, Dr. Basil Rigas, Dr. John Sninsky, Dr. Mark Sobel, Dr. Steven Sommer, and Dr. Thierry Soussi). Please see letters of collaboration in the overview section of this program project grant.

F. TIMETABLE

General: The various aims of this research will be pursued in parallel, with successful results in one arena providing impetus for progress on other aspects. The focus of Project 5 is the chemical synthesis of zip code DNA and PNA sequences, ideally in spatially addressable arrays, on appropriately optimized solid supports. As requisite materials and/or structures become available, they will be tested in relatively short order by Core 2. The list below follows a combination of descriptions in "Specific Aims" and in "Experimental Design and Methods."

Task 1. Design and optimization of zip code/address duplexes.

a. Synthesis of zip code sequences and their complements as DNA, followed by solution annealing studies. Months 1 to 12.
b. Synthesis of zip code sequences and their complements as PNA, followed by solution annealing studies. Months 7 to 24.
c. Synthesis of 5-propynyl-U monomers (Figures 3 and 4), synthetic incorporation into DNA and PNA, and solution annealing studies. Months 12 to 36.

Task 2. Development and evaluation of solid support materials compatible with chemical synthesis of DNA oligonucleotides and PNA oligomers, and compatible with subsequent hybridization reactions.

a. Studies with commercially available membranes and literature methods for immobilization of end-group modified DNA and PNA. Months 1 to 30.
b. Studies with "shaved" beads, which will be used for solid-phase synthesis of DNA. Months 1 to 18.
c. Modifications in our laboratories of surfaces, beads, or membranes with hydrophilic spacers such as heterobifunctional polyethylene glycol (PEG) and/or carbohydrates (see Figure 6 and accompanying discussion), and further studies. Months 12 to 48.
d. Development of novel chemistry for covalent immobilization of synthetic DNA or PNA (Figure 7). Months 6 to 30.
e. Preparation and segment condensation of protected PNA tetramer building blocks (including optimization of protection scheme and coupling conditions), to build up 24-mer complementary zip code sequences which will be released into solution following chain assembly. Months 12 to 48.
f. Segment condensation using protected PNA tetramer building blocks to build up 24-mer complementary zip code sequences which will be deprotected but retained on suitable solid supports for subsequent hybridization reactions. Months 36 to 60.

Task 3. Establishment of methodology for synthesis of spatially addressable arrays of DNA oligonucleotides and PNA oligomers.

a. Adaptation of commercially available membranes and literature methods for immobilization of end-group modified DNA and PNA, in tandem with spot methods and/or masking technology, to prepare and test relatively small arrays. Months 12 to 48.
b. Application of additional advances from Task 2 towards generation of spatially addressable arrays (e.g., "gluing" of "shaved" beads to solid surfaces, direct masking/segment condensation on solid supports as outlined in Figures 9 and 10). Months 12 to 60.

G. HUMAN SUBJECTS / VERTEBRATE ANIMALS: Not applicable

I. CONSULTANTS/COLLABORATORS: Project 5 represents a collaboration of Dr. George Barany (Principal Investigator), University of Minnesota, Dr. Robert Hammer, Louisiana State University, and Dr. Francis Barany, Cornell University, Medical College. In addition, we have excellent connections with leading industrial laboratories that are at the forefront of developing and commercializing methodologies for preparation of PEG-PS supports and functionalized membranes, PNA synthesis, and oligopolymer array construction. Specifically, we are collaborating with the team at Millipore led by Dr. James Coull, and a start-up company named Siris that is headed by Dr. Ronald Cook (see supporting letters). Recently, Dr. Michael Egholm, first author of several of the seminal papers on PNA, joined Millipore as a research chemist. Dr. Derek Hudson, a long-time collaborator of Dr. George Barany (several joint publications), is currently at Siris. Letters and Biographical Sketches for collaborators are attached in the overview section of this program project grant.

J. CONSORTIUM/CONTRACTUAL ARRANGEMENTS: Please see following page.

CORNELL UNIVERSITY MEDICAL COLLEGE
DEPARTMENT OF MICROBIOLOGY

1300 YORK AVENUE, Box 62
NEW YORK, N.Y. 10021
Telephone: (212) 746-6505
Fax: (212) 746-8587

STATEMENT OF INTENT TO ESTABLISH A CONSORTIUM AGREEMENT

Date: January 26, 1994

Grant Number: P01-

P-01 Application Title: PROGRAM PROJECT: NEW METHODS FOR CANCER DETECTION

Project # 5; DESIGN AND SYNTHESIS OF DNA AND PNA ARRAYS.

Proposed Project Period: Year 01; 12/01/94- 11/30/95

"The appropriate programmatic and administrative personnel of each institution involved in this grant application are aware of the NIH consortium grant policy and will establish the necessary inter-institutional agreement(s) consistent with that policy."

CORNELL UNIVERSITY MEDICAL COLLEGE, NEW YORK, NY

(Applicant Institution)

(name) (date)
Principal Investigator:

FRANCIS BARANY, Ph.D.

(name) (date)
Official Authorized to Sign for Institution

GREGORY W. SISKIND, M.D.
ASSOCIATE DEAN

UNIVERSITY OF MINNESOTA MINNEAPOLIS, MN.

(Consortium Institution)

(name) (date)
Co-Investigator:

GEORGE BARANY, Ph.D.

(name) (date)
Official Authorized to Sign for Institution

K. LITERATURE CITED.

1. Merrifield, R.B., *Solid-phase peptide synthesis.* J. Am. Chem. Soc., 1963. **85**: p. 899-904.

2. Barany, G. and R.B. Merrifield, *Solid-Phase Peptide Synthesis,* in *The Peptides,* E. Gross and J. Meienhofer, Editor. 1979, Academic Press: New York. p. 1-284.

3. Merrifield, R.B., *Solid phase synthesis.* Science, 1986. **232**: p. 341-347.

4. Barany, G., N. Kneib-Cordonier, and D.G. Mullen, *Solid-phase peptide synthesis: A silver anniversary report.* Int. J. Peptide Protein Res., 1987. **30**: p. 705-739.

5. Kent, S.B.H., *Chemical synthesis of peptides and proteins.* Ann. Rev. Biochem., 1988. **57**: p. 957-989.

6. Atherton, E. and R.C. Sheppard, *Solid Phase Peptide Synthesis: A Practical Approach.* 1989, Oxford: IRL Press.

7. Atherton, E., D.L.J. Clive, and R.C. Sheppard, *Polyamide supports for polypeptide synthesis.* J. Am. Chem. Soc., 1975. **97**: p. 6584-6585.

8. Smith, C.W., G.L. Stahl, and R. Walter, *Poly-N-acrylylpyrrolidine. A new resin in peptide chemistry.* Int. J. Peptide Protein Res., 1979. **13**: p. 109-112.

9. Atherton, E., *et al.*, *A physically supported gel polymer for low pressure, continuous flow solid phase reactions: Application to solid phase peptide synthesis.* J. Chem. Soc., Chem. Commun., 1981. : p. 1151-1152.

10. Arshady, R., *et al.*, *Peptide synthesis part 1: Preparation and use of polar supports based on poly(dimethylacrylamide).* J. Chem. Soc. Perkin Trans. I, 1981. p. 529-537.

11. Small, P.W. and D.C. Sherrington, *Design and application of a new rigid support for high efficiency continuous-flow peptide synthesis.* J. Chem. Soc., Chem. Commun., 1989. : p. 1589-1591.

12. Kanda, P., R.C. Kennedy, and J.T. Sparrow, *Synthesis of polyamide supports for use in peptide synthesis and as peptide-resin conjugates for antibody production.* Int. J. Peptide Protein Res., 1991. **38**: p. 385-391.

13. Bernatowicz, M.S., *et al.*, *Recent developments in solid phase peptide synthesis using the 9-fluorenylmethyloxycarbonyl (Fmoc) protecting group strategy,* in *Current Research in Protein Chemistry: Techniques, Structure, and Function,* J.J. Villafranca, Editor. 1990, Academic Press: San Diego.

14. Frank, R., *Spot-synthesis: An easy technique for the positionally addressable, parallel chemical synthesis on a membrane support.* Tetrahedron, 1992. **48**: p. 9217-9232.

15. Frank, R. and R. Döring, *Simultaneous multiple peptide synthesis under continuous flow conditions on cellulose paper discs as segmental solid supports.* Tetrahedron, 1988. **44**: p. 6031-6040.

16. Eichler, J., M. Beyermann, and M. Bienert, *Application of cellulose paper as support material in simultaneous solid phase peptide synthesis.* Collect. Czech. Chem. Commun., 1989. **54**: p. 1746-1752.

17. Lebl, M. and J. Eichler, *Simulation of continuous solid phase synthesis: Synthesis of methionine enkephalin and its analogs.* Peptide Res., 1989. **2**: p. 297-300.

18. Eichler, J., *et al.*, *Evaluation of cotton as a carrier for solid-phase peptide synthesis.* Peptide Res., 1991. **4**: p. 296-307.

19. Büttner, K., H. Zahn, and W.H. Fischer, *Rapid solid phase peptide synthesis on a controlled pore glass support,* in *Peptides-Chemistry and Biology: Proceedings of the Tenth American Peptide Symposium,* G.R. Marshall, Editor. 1988, Escom Science Publishers: Leiden, The Netherlands. p. 210-211.

20. Tregear, G.W., *Graft copolymers as insoluble supports in peptide synthesis,* in *Chemistry and Biology of Peptides,* J. Meienhofer, Editor. 1972, Ann Arbor Sci. Publ.: Ann Arbor, MI. p. 175-178.

21. Bayer, E. and W. Rapp, *Polystyrene-immobilized PEG chains: Dynamics and application in peptide synthesis, immunology, and chromatography,* in *Poly(Ethylene Glycol) Chemistry: Biotechnical and Biomedical Applications,* J.M. Harris, Editor. 1992, Plenum Press: New York. p. 325-345.

22. Barany, G., *et al.*, *Novel polyethylene glycol-polystyrene (PEG-PS) graft supports for solid-phase peptide synthesis*. ed. C.H. Schneider and A.N. Eberle. 1993, Leiden, The Netherlands: Escom Science Publishers. 267-268.

23. Zalipsky, S., *et al.*, *Preparation and applications of polyethylene glycol-polystyrene graft resin supports for solid-phase peptide synthesis.* Reactive Polymers, 1994. **in press.**

24. Jung, G. and A.G. Beck-Sickinger, *Multiple peptide synthesis methods and their applications.* Angew. Chem. Int. Ed. Engl., 1992. **31**: p. 367-383.

25. Geysen, H.M., R.H. Meloen, and S.J. Barteling, *Use of peptide synthesis to probe viral antigens for epitopes to a resolution of a single amino acid.* Proc. Natl. Acad. Sci. U. S. A., 1984. **82**: p. 3998-4002.

26. Fodor, S.P.A., *et al.*, *Light-directed, spatially addressable parallel chemical synthesis.* Science, 1991. **251**: p. 767-773.

27. Cass, R., *et al.*, *Pilot, A new peptide lead optimization technique and its application as a general library method,* in *Peptides - Chemistry, Structure and Biology: Proceedings of the Thirteenth American Peptide Symposium,* R.S. Hodges and J.A. Smith, Editor. 1994, Escom: Leiden, The Netherlands.

28. Lam, K.S., *et al.*, *A new type of synthetic peptide library for identifying ligand-binding activity.* Nature, 1991. **354**: p. 82-84.

29. Caruthers, M.H., *Chemical synthesis of DNA and DNA analogues.* Acc. Chem. Res., 1991. **24**: p. 278-284.

30. Garegg, P.J., *et al.*, *Nucleoside H-phosphonates. III. Chemical synthesis of oligodeoxyribonucleotides by the hydrogen phosphonate approach.* Tetrahedron Lett., 1986. **27**: p. 4051-4057.

31. Froehler, B.C., P.G. Ng, and M.D. Matteucci, *Synthesis of DNA via deoxynucleoside H-phosphonate intermediates.* Nucleic Acids Res., 1986. **14**: p. 5399-5407.

32. Uhlmann, A. and A. Peyman, *Antisense oligonucleotides: A new therapeutic principle.* Chem. Rev., 1990. **90**: p. 543-584.

33. Southern, E.M., U. Maskos, and J.K. Elder, *Analyzing and comparing nucleic acid sequences by hybridization to arrays of oligonucleotides: Evaluation using experimental models.* Genomics, 1992. **13**: p. 1008-1017.

34. Maskos, U. and E.M. Southern, *Oligonucleotide hybridisations on glass supports: a novel linker for oligonucleotides synthesised in situ.* Nucleic Acids Research, 1992. **20**: p. 1679-1684.

35. Maskos, U. and E.M. Southern, *Parallel analysis of oligodeoxyribonucleotide (oligonucleotide) interactions. I. Analysis of factors influencing oligonucleotide duplex formation.* Nucleic Acids Res., 1992. **20**: p. 1675-1678.

36. Maskos, U. and E.M. Southern, *A study of oligonucleotide reassociation using large arrays of oligonucleotides synthesised on a glass support.* Nucleic Acids Res., 1993. **21**: p. 4663-4669.

37. Fodor, S.P.A., *et al.*, *Multiplexed biochemical assays with biological chips.* Nature, 1993. **364**: p. 555-556.

38. Khrapko, K.R., *et al.*, *A method for DNA sequencing by hybridization with oligonucleotide matrix.* J. DNA Seq. Map., 1991. **1**: p. 375-388.

39. Van Ness, J., *et al.*, *A versatile solid support system for oligodeoxynucleoside probe-based hybridization assays.* Nucleic Acids Res., 1991. **19**: p. 3345-3350.

40. Zhang, Y., *et al.*, *Single-base mutational analysis of cancer and genetic diseases using membrane bound modified oligonucleotides.* Nucleic Acids Res., 1991. **19**: p. 3929-3933.

41. Egholm, M., *et al.*, *Peptide nucleic acids (PNA). Oligonucleotide analogues with an achiral peptide backbone.* J. Am. Chem. Soc., 1992. **114**: p. 1895-1897.

42. Egholm, M., *et al.*, *Recognition of guanine and adenine in DNA by cytosine and thymine containing peptide nucleic acids (PNA).* J. Am. Chem. Soc., 1992. **114**: p. 9677-9678.

43. Egholm, M., *et al.*, *Peptide nucleic acids containing adenine or guanine recognize thymine and cytosine in complementary DNA sequences.* J. Chem. Soc., Chem. Commun., 1993. : p. 800-801.

44. Egholm, M., *et al.*, *PNA hybridizes to complementary oligonucleotides obeying the Watson-Crick hydrogen-bonding rules.* Nature, 1993. **365**: p. 566-568.

45. Christensen, L., *et al.*, *Improved synthesis, purification, and characterization of PNA oligomers,* in *Innovation and Perspectives in Solid Phase Synthesis and Complementary Technologies: Biological and Biomedical Applications - 1994,* R. Epton, Editor. 1994, SPCC (UK) Limited:

46. Vagner, J., *et al.*, *Novel methodology for differentiation of "surface" and "interior" areas of polyoxyethylene-polystyrene (POE-PS) supports: Application to library screening procedures,* in *Innovation and Perspectives in Solid Phase Synthesis and Complementary Technologies: Biological and Biomedical Applications - 1994,* R. Epton, Editor. 1994, SPCC (UK) Limited:

47. Fields, G.B., Z. Tian, and G. Barany, *Principles and Practice of Solid-Phase Peptide Synthesis,* in *Synthetic Peptides: A User's Guide,* G. Grant, Editor. 1992, W.H. Freeman and Co.: New York. p. 77-183.

48. Barany, G. and F. Albericio, *Recent progress on handles and supports for solid-phase peptide synthesis,* in *Peptides-Chemistry, Structure and Biology: Proceedings of the Thirteenth American Peptide Symposium,* R.S. Hodges and J.A. Smith, Editor. 1994, Escom Science Publishers: Leiden, The Netherlands. p. in press.

49. Albericio, F., *et al.*, *Preparation and application of the 5-(4-(9-fluorenylmethyloxycarbonyl)aminomethyl-3,5-dimethoxyphenoxy)valeric acid (PAL) handle for the solid-phase synthesis of C-terminal peptide amides under mild conditions.* J. Org. Chem., 1990. **55**: p. 3730-3743.

50. Albericio, F. and G. Barany, *Hypersensitive acid-labile (HAL) tris(alkoxy)benzyl ester anchoring for solid-phase synthesis of protected peptide segments.* Tetrahedron Lett., 1991. **32**: p. 1015-1018.

51. Froehler, B.C., *et al.*, *Oligonucleotides containing C-5 propyne analogs of 2'-deoxyuridine and 2'-deoxycytidine.* Tetrahedron Lett., 1992. **33**: p. 5307-5310.

52. Harris , J.M., ed. *Poly(Ethylene Glycol) Chemistry: Biotechnical and Biomedical Applications.* 1992, Plenum Press: New York.

53. Lofas, S. and B. Johnsson, *A novel hydrogel matrix on gold surface plasma resonance sensors for fast and efficient covalent immobilization of ligands.* J. Chem. Soc., Chem. Commun., 1990. : p. 1526-1528.

54. Goodchild, J., *Conjugates of oligonucleotides and modified oligonucleotides: A review of their synthesis and properties.* Bioconjugate Chem., 1990. **1**: p. 165-187.

55. Kneib-Cordonier, N., F. Albericio, and G. Barany, *Orthogonal solid-phase synthesis of human gastrin-I under mild conditions.* Int. J. Peptide Protein Res., 1990. **35**: p. 527-538.

56. Barany, G. and F. Albericio, *A three-dimensional orthogonal protection scheme for solid-phase peptide synthesis under mild conditions.* J. Am. Chem. Soc., 1985. **107**: p. 4936-4942.

Appendix 2

(To Declaration of Francis Barany under 37 CFR § 1.608(b))

DRAFT REVIEW REPORT

FRANCIS BARANY, PH.D.

1 P01 CA65930-01

"NEW METHODS FOR CANCER DETECTION"

CORNELL UNIVERSITY MEDICAL COLLEGE
NEW YORK, NEW YORK

Committee/Date: Special Review Subcommittee C/July 20-22, 1994

Review Meeting/Date Held: Site Visit/May 31-June 2, 1994

Special Notes: None

OUTSIDE OPINION OBTAINED
Administrative Note: Page 19.
Resume: Funds are requested for this application that describes a series of exciting technological advances for detecting high sensitivity mutations that could have a profound impact on the diagnosis, and eventually therapy, of human cancer. The application describes several clinical situations where high resolution detection of mutations in general, and base substitution mutations in particular, could have a major impact on the early detection and diagnosis of cancer, prognostication based on micrometastases at the time of diagnosis, and the early detection of recurrence. The strengths include the potential of the proposed studies, and the high capability of the investigators to carrying out the proposed studies. The weaknesses include the overly ambitious nature of the application, a lack of experience of the investigators in cancer related work, a lack of the proposed clinical correlations, a lack of biostatistical consultation in the design of the studies, and the serious flaws involved in the proposed pilot studies of the clinical utility of the approach. If considered solely as an exercise in technology enhancement this proposal could have ranked in the outstanding range. This program project is rated at a very good to excellent level of merit.

TABLE OF CONTENTS

OVERALL DESCRIPTION (Applicant's description)

The long range objective of this proposal is to develop sensitive and specific approaches to the detection and stimulanteous identification of cancer-related, genetic alterations. Mutations and genetic aberrations have been implicated, at various steps, in the etiology and biology of tumors. Inherited mutations account for the predisposition to cancer in some families. Somatic mutations in tumor suppressor genes, oncogene amplification and viral DNA sequences have been found in cancers as well. However, the clinical use of these discoveries and research into their clinical significance has been slowed by the laborious process by which they are detected. To apply these discoveries and explore the interactions of multiple genetic alterations, we urgently need a new technology, which is capable of being automated and has the power to detect any of a vast number of mutations.

In response to the urgent need for new methods of mutation detection, we have assembled a team of investigators whose expertise will be directed toward innovative solutions to this problem. The collaborative nature of the scientific and organizational infrastructure will facilitate the attainment of the projects' specific aims and objectives.

The specific aims of the five projects in this program project are to: (i) develop a multiplex polymerase chain reaction/ligase detection reaction (PCR/LDR) system for the detection of inherited mutations in germline DNA and somatic mutations in tumors; (ii) develop a ligase detection reaction/polymerase chain reaction (LDR/PCR) system for detecting gene amplifications and deletions in tumors; (iii) develop a PCR/restriction /LDR (PCR/RE/LDR) system for detecting and identifying mutations in rare cancer cells at a sensitivity of 1 in 106 or 1 in 107 by removing normal DNA sequences and selectively amplifying cancer mutations; (iv) design and synthesize nucleotide analogues for converting specific DNA sequences into restriction endonuclease recognition sites for PCR/RE/LDR mutation detection; (v) engineer a thermostable ligase with greater fidelity to enhance LDR and LCR specificity; (vi) design and synthesis oligonucleotide or peptide nucleic acid (PNA) addressable arrays for the simultaneous detection of multiplex LDR and LCR products; and (vii) explore the ability of these technologies to further our understanding and clinical management of lung, colon, breast and cervical cancers.

OVERALL CRITIQUE

The goal of this program project is to develop sensitive PCR/LCR and very sensitive PCR/RE/LCR base substitution mutation tests which could be used to screen, in multiplex, large numbers of tumors for important mutations related to cancer (PCR/LCR) and to screen large numbers of cells for the presence of these mutations (PCR/RE/LCR). The technology development effort in Project 1, together with Project 2 could be outstanding, although the separation of the two projects appears artificial.

The enthusiasm for Project 1, led by Dr. Wilson, is tempered by an inadequate discussion of potential false positive results. More important, is the weakness of the proposed clinical correlations. The experimental plan for these correlations is diffuse and poorly described. The screening of 90 tumors by PCR/LCR does not permit sufficient statistical power to allow for any correlations with clinical outcome and no acceptable plan is put forward to achieve this end. Given that the diagnosis of recurrence in lung and colon cancer requires examining known sites of metastases (e.g. liver, bone) and the inability of current treatment to significantly impact on survival after documentation of recurrence, it is not clear how PCR/RE/LCR would permit early detection of metastases in a useful way. A focus on the prognostic significance of micrometastatic disease at the time of diagnosis or on the detection of known mutations in cells not yet cytologically "malignant" would have far greater utility. It is not clear what the second clinical site in Denver adds to the overall project. The feasibility of the proposed technologies could be tested adequately on the samples already located in New York. Overall, this project is rated as very good.

Project 2, headed by Dr. Francis Barany, is considered to be the stronger of the two clinical projects and stronger enthusiasm is expressed for the application of the proposed technology to the problems inherent in working with clinically

heterogenous tissue such as that found in human breast cancer. As in Project 1, there is considerably less enthusiasm for the clinical-correlative studies. The investigators did not present a convincing case that they understood the extent of the clinical base of materials available and whether they knew how to apply their assays even in a preliminary way. The project would have benefitted from the up-front collaboration of a biostatistician. The studies proposed for cervical cancer, although interesting, are not considered to add anything to the project and could be dropped. Overall, the project is rated as excellent.

Project 3 is directed by Dr. Donald Bergstrom. The proposed design and synthesis of candidate convertides and universal bases for development of new and useful oligonucleotide diagnostics represent a largely empirical, albeit scientifically sound and potentially highly significant, effort. The structures and synthetic routes appear reasonable. Promising results have already been obtained and it appears that the proposed iterative "synthesis-testing-synthesis" scheme is feasible. In view of the challenging synthetic goals and the lack of a convincing rationale for the propony1 oligonucleotides, it is recommended that this latter element be reconsidered. The potential for this project to have a significant impact on nucleic acid based technology leads to an overall rating of outstanding.

Each of the three specific aims proposed in Project 4, led by Dr. Francis Barany, has the potential to generate important information about the Tth DNA ligase in particular, and DNA ligases in general. With respect to the applicability to the rest of the program project, the proposed determination of the structure of the ligase is considered to be the weakest part of the project. The probability that the structure will actually be solved early enough in the project to be applicable to the other projects is considered remote, although the work itself is of major importance to the field. The ability of an improved ligase to enhance the sensitivity of the LCR/LDR assays, and hence their potential diagnostic value leads to an overall rating of excellent for this project.

In order to carry out the large-scale screening of mutations, the Project Leader, Dr. George Barany (Project 5), proposes to develop spatially addressable arrays of oligonucleotides or peptide oligonucleotide analogs. Ligase reaction products will be constructed with fluorescent groups and will bear specific "zip code" tails. The tails will be selectively captured by complementary zip code probes immobilized in the array. Each zip code will map for a specific, known, genetic mutation. Serious concern is expressed over the ability of this workplan to actually deliver a working array to the other projects. The project lacks adequate planning in hybridization, signal to noise ratios, array fabrication expertise and instrumentation development. On the other hand, the concept of zip code sequences and the potentially elegant contributions to solid-phase chemistry by the project leader are considered strong assets. Also recognized, is the broad applicability of the technology to areas outside of cancer diagnostics. Overall, the project is rated as very good to good.

Core A, headed by Dr. Niel Hackett, proposes to: (1) create and maintain a relational database from existing clinical databases in New York and Denver; (2) perform statistical correlations of clinical outcomes with laboratory studies; (3) develop software for the selection of appropriate primers and zip codes; and (4) continue support of both instrumentation and connectivity. All four elements of this facility are considered essential; however, the experience of the Core Leader is considered

sufficient only for elements 3 and 4. There is no demonstrable expertise in the acquisition or analysis of clinical data and this is a major weakness of the overall application. Thus a merit rating of this Core is at acceptable level of merit.

Core B (Dr. Francis Barany) is recognized as an essential part of this program project. The reviewers consider that the primary function of this core should be to provide oligonucleotide reagents to the rest of the program, and to a lesser extent, provide a testing service for polymerase fidelity and efficiency of nucleotide conversions. This Core would benefit from additional expertise in the areas of fluorescence instrumentation and the implementation of robotics. Overall, this core is rated as excellent.

Core C, directed by Dr. Francis Barany addresses the need to coordinate the interaction of the applicant organization with six separate consortium institutions. The projects are proposed to operate largely independently, with this core focusing on communication and reporting requirements. Dr. Barany has demonstrated an ability to develop substantive scientific collaborations, but evidence of administrative systems to promote and monitor these interactions is lacking. Dr. Buck, the administrative co-investigator, has been recruited on a part-time basis from the Strang Clinic to address these systems but his contribution to-date is not evident. An impressive panel of external advisors has been recruited, but the means to internalize their advice and maximize the efficacy of an annual review have not been addressed. A vast number of letters of support are provided that propose specific studies unrelated to this grant while providing strong support for the overall concept. The overall rating for this core is good to acceptable.

Program as an Integrated Effort: Although this program is well integrated and has synergy as far as its goals are concerned, the diverse locations of the laboratories along with a lack of a plan to hold the frequent meetings between the project leaders, on the part of Dr. Francis Barany, hinder the interactions between the investigators. However, formulation of a plan by the Principal Investigator to hold bimonthly meetings could facilitate such efficient interactions.

PRINCIPAL INVESTIGATOR

Dr. Barany received his Ph.D. degree in Microbiology from The Rockefeller University in 1981 where he also spent a year for a postdoctoral training. He then moved to Dr. Hamilton Smith's laboratory at The Johns Hopkins University for a three year postdoctoral training in Molecular Biology. He returned to New York in 1985 to assume a faculty position at the Cornell University Medical School where he is currently an Associate Professor of Microbiology. He is also an Adjunct Associate Professor at The Rockefeller University. His scientific productivity is reflected by his numerous publications in highly competitive refereed journals. However, Dr. Barany clearly expresses a lack of experience in dealing with the problems posed by human cancer and describes a series of existing and potential collaborations to overcome this issue. He does not demonstrate previous experience in managing a project of this magnitude but his ability to pull together and provide leadership to this program has led the reviewers to believe that he is well qualified to serve as the Principal Investigator.

SUPPORT TO BE NEGOTIATED FOR REPLACEMENT

Investigator	Grant Number
F. Barany	USAMRDC Pending ("Multiplex Detection of Point Mutations, Amplifications, and Deletions in Breast Cancer")
M. Lubin	USAMRDC Pending ("Multiplex Detection of Point Mutations, Amplifications, and Deletions in Breast Cancer")

INDIVIDUAL PROJECTS AND CORES

Project 1: Genetic Markers of Lung and Colon Cancer
(Vincent Wilson, Ph.D.)

Description: (Applicant's description) In the last decade mutations in many oncogenes and tumor suppressor genes have been described in cancers. This knowledge, however, has not significantly changed the care of cancer patients. Do cancer mutations predict the behavior of tumors? To correlate mutations with clinical outcomes we need robust methods to identify many possible mutations. Can the early spread of cancer be determined by finding the mutations of cancer cell in the bone marrow? To detect micrometastases or early cancers we must be able to detect a few cancer cells out of many normal cells.

To achieve these capabilities we have devised two technologies: polymerase chain reaction/ligase chain reaction (PCR/LCR) to survey tumors for a wide number of mutations simultaneously; and polymerase chain reaction/restriction endonuclease digestion/ligase chain reaction (PCR/RE/LCR) for detecting a few cancer cells out of many normal cells. When they are fully developed PCR/LCR should be able to detect tens to hundreds of mutations at a sensitivity of one in 102 or 103. PCR/RE/LCR has already detected one mutation-bearing cell out of 107 normal cells.

To demonstrate the feasibility of these methods our specific aims are to: (i) Develop a PCR/LCR multi-gene, multi-mutation detection system to simultaneously identify mutations in three condons of the k-ras oncogene and nine condons of the p53 tumor suppressor gene. Approximately half of colon cancers have these k-ras mutations. About 15 percent of lung tumors and about 21 percent of colon cancers have one of these nine p53 mutations. Using PCR/LCR to identify these mutations we will investigate 40 colon and 50 lung tumors; (ii) Refine PCR/RE/LCR to detect the above p53 mutations at sensitivities of one in 107. We will first use PCR/RE/LCR to determine the natural background mutation rate in non-cancerous tissues. Then, for patients whose tumors had detectable p53 mutations, we will use PCR/RE/LCR to investigate lymph nodes, blood and bone marrow specimens for micrometastases.

Critique: The project leader proposes to amplify genomic fragments from cancer-related genes using PCR and to rely on the allele specificity of the ligase chain reaction (hence PCR/LCR) to detect specific point mutations. Additional selection

is proposed to be added during the PCR amplification by way of restriction endonuclease cleavage of the wild type sequences (PCR/RE/LCR).

The project leader makes a compelling case that LCR (or LDR in Project 2) is a good method for detecting mutations in clonal tumor populations or in tissue samples in which the tumor cells represent 0.1-1 percent of the cells. Ease of multiplexing was put forward as the major advantage of LCR over allele-specific PCR. Nevertheless, unlike LDR in project 2, no preliminary data is presented for PCR/LCR with N>2. For the work scheduled for years 1-3, the project leader has selected mutations in 3 codons of k-ras (20 mutations) which will detect mutations in about 50 percent of colon and an undisclosed number of lung cancers, and mutations in 9 codons of p53 found in 15 percent of lung and 21 percent of colon cancers. PCR/LCR would be applied to 40 colon and 50 lung tumors from a more than adequate supply of available specimens. It is not clear how these specimens would be selected. No power calculations are given to assess the effort needed to extract clinically relevant data (e.g. prognosis) from this type of survey. It is also disappointing that no calculation is given as to the effort necessary to develop a multiplex PCR/LCR system capable of detecting mutations in a large fraction (e.g. 95 percent) of either lung or colon cancers, a prerequisite to screening or other applications of PCR/RE/LCR. Also, the discussion of the use of readily available fixed specimens is inadequate.

The project leader has also made a compelling case that PCR/RE/LCR could be developed in years 2-5 into a highly sensitive method for detecting mutations in tissue samples where tumor cells represented approximately 1/10 6 cells. The competing technology of PCR/RE/allele-specific PCR is not compared, although the previous advantage of potentially easy multiplexing in PCR/LCR would be lost in PCR/RE/LCR. No explanation was proffered as to why dilution of mutant plasmid into wild type plasmid led to decreased signal intensity (p. 218 of the application) whereas dilution of 1-10 cells containing mutant DNA into wild type cells led to constant signal intensity (p. 220). Furthermore, the project leader stated that experiments with 1 cell are routinely positive. Given the required sensitivity and discrimination of these experiments, data showing detection as a function of PCR cycle number for the various amplifications would have been reassuring. Although he is an expert at applying PCR technology, the discussion of contamination is inadequate for the task at hand. PCR/RE/LCR is equivalent to sperm typing, where heroic efforts have been needed to eliminate false positives from amplification product carry over.

A discussion of potential false positives due to DNA polymerase errors is included, but the assumed DNA polymerase error rate is at least an order of magnitude lower than the cited literature. Ligase may be contributing of the order of 100-fold to the discrimination. Nevertheless, the results (Fig. 4 & 6) for PCR/RE/LCR at MspI sites, which contain only dG and dC and which are copied by Taq polymerase with greatest fidelity, might suggest that the same degree of PCR/RE discrimination could not be achieved for the majority of mutations.

The proposal to examine various clinical samples for micrometastases, except as applied to establishing initial staging, appears to be misdirected to tissues unrelated to metastatic spread.

Intrinsic Scientific Merit Score: 221

Personnel:

Name: Vincent Wilson
Degree/Discipline/Date: Ph.D., pharmacology & toxicology, 1980
Role/Percent Effort: Project Leader, 20 percent
Qualifications/Experience: Dr. Wilson was a Senior Staff Fellow in the Laboratory of Human Carcinogenesis, NCI (1982-1988) where he published extensively with Dr. Curtis Harris on DNA adducts. His recent and independent publication record has not been particularly impressive. The work on which this study is based derives from his SPORE pilot project and his collaboration with the applicant. They have developed a highly sensitive method for detecting point mutations in one cell in millions, which if generalized could be important for cancer biology. Dr. Barany credits Dr. Wilson with the original concept of the application.

Assessment in designated role: highly qualified

Name: Leonid L. Reznikov
Degree/Discipline/Date: M.D., andrology and uroloncology, 1986; Ph.D., laser medicine, 1990
Role/Percent Effort: Postdoctoral fellow, 100 percent
Qualifications/Experience: His recent publications, in Russian, have all been in urology. He has no background in molecular biology.

Assessment in designated role: qualified.

Budget: The modest budget is approved as requested.

Assessment: Level of merit; very good

Project 2: Genetic Markers of Breast and Cervical Cancer
(Francis Barany, Ph.D.)

Description: (Applicant's description) To improve cancer care, researchers and clinicians need robust methods of identifying genetic alterations in cancers. There are three important challenges that need to be met: (i) the detection of many possible point mutations in tumors; (ii) the quantification of gene amplifications and deletions in tumors; and (iii) the detection of rare cancer cells against a background of normal cells. Researchers need these capabilities to be able to correlate multiple genetic alterations with clinical outcomes, identify new cancer-related genetic loci, and detect early cancer recurrence and premalignant cell.

To accomplish this, we will develop: (i) a multiplex polymerase chain reaction/ligase detection reaction (PCR/LDR) system to detect many possible point mutations in cancers; (ii) ad multiplex ligase detection reaction/polymerase chain reaction (LDR/LCR) system to quantify gene amplifications and deletions in tumors; and (iii) a polymerase chain reaction/restriction endonuclease/ligase detection reaction (PCR/RE/LDR) to identify 1 cancer cell in 106 normal cells.

Some issues in cancer will be explored. Specifically we will: (i) Expand PCR/LDR to detect 24-40 point mutations, (63% to 79% of p53 gene mutations) in breast tumors. PCR/LDR will also be used to detect high risk human papillomavirus (HPV) in cervical

lavages or biopsies. (ii) Use "zip code" primers to proportionally PCR amplify and quantify LDR products of genes deleted or amplified in tumors. This should allow us to simultaneously detect HER-2/neu and int-2 gene amplifications, as well as p53 gene deletions in breast tumors. (iii) Refine PCR/RE/LDR for detecting five p53 mutations by selectively amplifying mutated DNA while removing wild-type products by TaqI restriction endonuclease cleavage. We will then characterize p53 gene mutations and deletions, HER-2/neu amplifications, and int-2 amplifications in 100-200 frozen breast tumors and corresponding fixed specimens. We will be use PCR/RE/LDR to look for micrometastases in the bone marrow and lymph nodes of patients whose tumors had one of the five specific p53 mutations. Ultimately, this sensitivity method may identify early relapses or primary tumors by detecting circulating cancer cells in the blood.

Critique: In Specific Aim 1, a multiplex PCR/LDR system will be developed to detect mutations in the p53 gene and to detect high risk HPV strains in clinical samples. The strength of the project is that rapid methods to detect single base pair mutations are much needed in the clinical arena, and the PCR/LDR and the PCR/RE/LDR methods may prove useful for this purpose. From a technical point of view, the boundaries between this project and Project 1 are artificial; this project is distinguished from Project 1 primarily by the fact that Project 1 concentrates on LCR methods while this project concentrates on LDR methods. The three specific aims entail PCR/LDR of tumor biopsies and Human Papilloma Virus (HPV), LDR/PCR for measuring gene amplification, and PCR/RE/LDR for detecting mutations at a sensitivity of 1 in 10^6 or 10^7. In contrast to project 1, the target genes are directed towards genes involved in breast cancer (and HPV infection associated with cervical cancer), with an overlap with project 1 in regard to detecting some p53 mutations. The project leader hopes to identify new correlations with prognosis and the mutations revealed.

The project leader has chosen to initially examine five different mutation sites within the p53 gene (constituting about 28% of reported p53 mutations) because these are known sites for TaqI conversion using PCR/RE/LDR. This specific aim proposes to develop the method for large screen testing using a pilot panel of 100 to 200 breast cancer cases from frozen or fixed specimens. However, no preliminary data was shown using a sample with a known mutation at one of these sites in either frozen or fixed breast cancer samples. Furthermore, it is unclear from the letters of collaboration, the nature of the corresponding follow-up clinical information on these patients. The letter of collaboration from Dr. Kovach states that 75 known p53 mutation sites-containing samples would be made available, but it is not clear which of these contain the five p53 mutations chosen for study. The letter from Dr. Summers does not state any sample information, but appears to be referring to the wrong grant application. However, it is apparent that the well-characterized tumor bank available through Dr. Osborne is a valuable resource which will be made available to the investigators.

This project refers to project 4 because improvements in the buffer conditions or ligase may be needed in order to increase the sensitivity of PCR/LDR to 1 cancer gene mutation in 10^2 to 10^3 normal cells. If one accepts that the need for further optimization of pH , salt concentration, incubation time and temperature that give maximum fidelity for Tth ligase in LDR assay, and thus LCR (see page 356 of project 4), then the experiments proposed seem premature. It is most important to optimize the signal before performing complex experiments.

Another issue is that the investigators have chosen the p53 gene because of the frequency with which mutations have been detected in tumors, and it is a valuable test system for developing the methods of PCR/LDR and PCR/RE/LDR. But it is questionable whether the investigators will be able to correlate their findings due to their current lack of integration with the clinical database and their lack of statistical power. Even very good prognostic markers with hazard ratios of 1.3 to 1.4 and rare mutations rates (for instance of 8 percent) would require 1000 to 1700 specimens to see differences in disease free survivals (in node-negative disease) after five years. Additionally, there is no discussion of statistical analysis in the project; a statistical collaborator would strengthen this project.

It is of concern that cellular heterogeneity inherent in breast cancer has not been adequately addressed; this may be problematic in PCR/LDR assays. It would probably be best to develop Specific Aim 3 first, and increase sensitivity, before screening large numbers of samples outlined in Aim 1.

The project leader has also included studies to detect high risk HPV strains in cervical carcinomas. It is felt that the inclusion of this study, when the rest of the project is devoted to the detection of breast cancer-specific changes, makes the project diffuse. It is suggested that this study be omitted.

The arguments for gene quantification as a method to look for ploidy changes are quite convincing. However, figure 7 in the application is a multiplex experiment in which the ratios should be 1:1 but they are not; this is not discussed by the Project leader. It is stated that "the prognostic significance of gene amplifications (e.g. HER-2, c-myc, and int-2) in breast cancer has not been clearly established..)". There are numerous studies demonstrating the clinical utility of HER-2 in node-positive disease, and the lack of utility of int-2 as a prognostic marker. It is not understood why int-2 has been chosen for study due to its low level amplification rate in breast cancer (less than 15 percent) and due to the fact that the protein product of int-2 is not expressed in breast cancer. Although the development of an assay for gene amplification is worthy of study, the development of it for genes such as the int-2 gene is both scientifically uninteresting, and a poor choice for study. However, the HER-2 gene is an appropriate choice for methodology development if HER-2 amplification status has been previously determined by conventional methodologies in these clinical samples.

In Specific Aim 3, PCR/RE/LDR will be developed to detect rare mutations with increased sensitivity for the eventual application of occult micrometastasis identification; this aim is undoubtedly the strength of the project. The PCR/RE/LDR strategy could prove to have a relatively high sensitivity in clinical samples. As the investigators point out, the level of error incorporation of Taq polymerase is unlikely to average over 1 in 10^{-6} cells, and furthermore, the ligase has a 50x to 500x selectivity against mismatches involving the 3' nucleotide of the ligation substrate. Preliminary data would have dispelled these doubts.

To generalize PCR/RE/LDR, the project Leader suggests conversion of mutation sites to include a restriction site by the use of primers with "convertide" nucleotides or nucleotide analogs. The example given is for transitions from CCGG to TCGA, the easiest of such conversions. The investigators propose that other harder conversions could use nucleotide analogs that permit the introduction of other bases in replication (project 3).

The investigators in Project 3 may have already found a universal "convertide" for this purpose. The other proposed convertides would not now be deemed necessary for this project. They would still be interesting as variants at the penultimate 3' base for increased discrimination. It should be noted that there is no reference to previous work regarding "converting" sequences to restriction sites using PCR. These include Hruban et al., Am. J. Path., 1993 and Mitsudomi et al., 1991. These alternative technologies should have been discussed.

From the point of view of error rate in polymerization or ligation, the examples shown generally represent atypically easy targets for demonstrating this strategy; the restriction site is 5'-CCGG (or 5'-TCGA with the T and A defined by the primers) and it is known that Taq polymerase has its lowest error rate for G:C base pairs. The mutations detected are usually transversions, for which the ligase has the maximum possible discrimination. This means that the example shown will not reflect the typical combination of A:T-containing restriction sites and transition mutations that will be encountered in a big survey. The problem of fidelity at some sites was acknowledged at the visit.

Once conversion of a site is contemplated there can be no multiplexing. One should also consider whether allele-specific ASO (similarly spiked to give quantitation) would be equally efficacious, or even more so, after iterated PCR/RE. Such a PCR step could take place under maximum fidelity conditions (such as single stranded binding protein and low nucleotide concentrations) which seems to match the fidelity of LCR. Specific reasons why this is not a good idea should be given.

The investigator recognizes that in many of the cases for which this methodology may be useful, the issue will not only be that of whether a mutation is present, but also how prevalent this mutation might be. To achieve this the reaction will be spiked with a primer containing a different 3' base than the one to be assayed and with a different product length. This strategy assumes that the chosen spike sequence will never occur in a tumor but is nevertheless likely to work in many cases.

Regardless of these concerns, a high degree of enthusiasm is expressed for this study, independent of the ability of the project leader to correlate the information gained with clinical parameters.

Intrinsic Scientific Merit Score: 183

Personnel:

Name: Francis Barany
Degree/Discipline/Date: Ph.D., microbiology, 1981
Role/Percent Effort: Principal Investigator, 15 percent
Qualifications/Experience: described earlier.

Assessment in designated role: highly qualified.

Name: Matthew B. Lubin
Degree/Discipline/Date: M.D., medicine, 1984
Role/Percent Effort: Co-Investigator, five percent

Qualifications/Experience: Internship and Residency, and Fellowship in Internal Medicine, and Medical Genetics, respectively (1987-1990); Clinical Instructor (1990-1993); Director of Medical Genetics, and Assistant Professor (1990 and 1993).

Assessment in designated role: well qualified.

Name: Darren Day
Degree/Discipline/Date: Ph.D., biochemistry, 1989
Role/Percent Effort: Research Associate, 100 percent
Qualifications/Experience: Postdoctoral Fellow at the University of Southampton, United Kingdom, and Auckland University, New Zealand (1990-1992).

Assessment in designated role: qualified.

Budget: The requested budget is approved.

Assessment: Level of merit; excellent.

Project 3: Design and Synthesis of Nucleotide Analogues
(Donald Bergstrom, Ph.D.)

Description: (Applicant's description) A high sensitivity mutation detection system must be able to detect changes in any gene sequence. To be able to accomplish this, wild type DNA sequence corresponding to a mutation needs to be converted to a restriction enzyme site so repeated PCR amplification followed by digestion with the restriction enzyme removes the normal sequence while selectively amplifying the mutant sequence. This amplification, known as PCR/RE/LDR aims to detect one cancer mutation in 106 normal cells.

The goal of this project is to design and synthesize nucleotide analogues which facilitate sequence conversion. "Convertides" are nucleoside analogues which pair to one or more of the natural bases in an initial primer hybridization. More importantly, convertides also function as a degenerate template allowing for insertion of different base during subsequent rounds of polymerase amplification. There are twelve possible nucleotide conversions which should be achieved.

To accomplish our goal we will work towards the following specific aims: (i) The synthesis of deoxyribonucleoside analogues to be used as convertides. Eight deoxyribonucleoside analogues, Q2, Q5, Q6, and Q9-Q13, have been previously described. We have already designed nine additional modified deoxyribonucleosides, Q1, Q3, Q4, Q7, Q8, and Q14-Q17. All 17 deoxyribonucleosides analogues will be synthesized in our laboratories. (ii) Preparation of dimethoxytrityl (DMT)-protect derivatives of all the convertides for incorporation into oligonucleotides. In the middle of an oligonucleotide, DMT-convertide phosporamidites will be used. At the 3' position this will be accomplished by attaching the 3'-hydroxyl of the protected convertide to a long chain alkyl amine-CPG support. (iii) Testing of convertides for use in the mutation detection techniques. Starting with Q2, Core B will test convertide oligonucleotides as a means of increasing the specificity of mutation detection and as universal bases for polymorphic sites. (iv) Synthesizing and incorporating 5-propynyluridine into DNA or PNA "zip codes". In addressable arrays, this will be tested for optimizing the Tm of the zip codes/complementary zip code duplexes (Project 5 and Core B).

Critique: The first Specific Aim of this project is to synthesize 17 deoxynucleoside analogues (Q1-Q17) to be incorporated into convertide oligonucleotide probes and primers; 8 analogues have been previously reported and 9 are apparently new. The underlying theory for the selection or design of each of these particular analogue structures seems reasonable. The H-bonding bonding schemes and tautomeric forms shown on p.321 of the application are also reasonable, with the exception of Q4 and Q8 in which oxygen has been replaced by sulfur. It would have been useful in each of these two less familiar cases to include supportive thermodynamic calculations. The same is true for several C-nucleosides, the duplex-forming properties of which are not clear-cut. On the other hand, the applicant's rationale regarding the selection or design of Q1-Q17 is supported by his preliminary studies of Q2 that include thermal melting (Tm) measurements to determine if Q2 is non-discriminatory in base-pairing to A, C, G, and T, as well as investigation of sequencing and PCR reactions. While the Tm resulted per se are can only be suggestive of non-discriminatory base-pairing of Q2, it is very important that an oligonucleotide with Q2 (but not with mismatches) was already shown by the applicant to be functional in primer extension in PCR and by T7 DNA polymerase even with Q2 located at the 3'-end of a primer, which is a critical locus for the proposed convertide mechanism.

The applicant's preliminary synthetic results for Q14 are likewise supportive. Known-target analogues Q12 and Q13, the synthesis of which are said to be in progress along with Q5 and Q9, have already been shown by others to be tolerated by Taq polymerase when at the 3' end of a primer. This gives this study further credence with regard to feasibility.

In summary, the structural targets proposed in the first Specific Aim represent novel and challenging design and synthesis efforts that are correctly viewed by the applicant as being an empirical screening process to find which candidate convertides will actually provide acceptable biochemical "read" and "write" kinetics. While this adds some uncertainty to the degree of success achievable with each of the 17 convertides presently proposed, there seems to be no other way to approach the problem at this stage, since even state-of-the-art molecular modeling can at best be used to assess relatively simple H-bonding schemes and duplex stability but none of the critical polymerase reactions.

Regarding synthesis, half of the targets are known and should therefore pose no significant difficulties. The new convertides involve largely conventional transformations with apparently reasonable literature precedent. These syntheses will involve a substantial amount of work but are otherwise unremarkable. Alternative synthetic schemes are said to be available, although not written in the application.

The second Specific Aim, which is the preparation of DMT-protected support-bound and phosphoramidite derivatives of the convertides, has been adequately addressed using relatively straightforward chemistry that should pose no significant problems for the applicant's team. Purification and analysis of these oligonucleotides is described only very briefly; however, again there is ample precedent for this type of post-synthetic work.

The third Specific Aim involves testing of convertides as a means of increasing the specificity of mutation detection and serving as functionally useful bases

for polymorphic sites, which is work carried out in program projects 1, 2, and 4. Promising convertides will be further studied by the applicant using conventional measurement of thermodynamic parameters.

The fourth Specific Aim concerns synthesis of 5-propynyl-dU by the Method of Froehler in order to obtain modified oligonucleotide with supposedly higher hybridization affinity and thus, in principle, obtain greater discrimination in the addressable arrays described in Project 5. A similar proposal was made for 5-propynyluracil PNA monomers with either Fmoc or Boc protection. This is an interesting idea in the case of dU based on Froehler's published initial work; however, the generality and magnitude of this effect in mixed-based sequences has not apparently been reported. There is apparently no experimental precedent that the 5-propynyl effect will apply to PNA-DNA heteroduplexes, nor is any supporting theoretical rationale developed by the applicant. Moreover, no already proven options such as 2'-O'alkyl oligonucleotides are even discussed. While this aspect of the work is far less critical to realization of the convertide concept, it nevertheless ought to be reconsidered.

In summary, this is an exciting project based upon the early work of others and preliminary promising results obtained by the applicant. It can contribute novel compounds for potentially promising and generally useful enhancements to hybridization-based detection and a justification schemes proposed in the program and possibly beyond. The synthesis of convertide-containing oligonucleotides by the present project is clearly critical for the success of the overall program. A specific task time-table is not developed indicating who is to make which convertides by specified projected dates for delivery to the other projects and core components; however, the synthetic work is distributed between Drs. Bergstrom and Hammer according to the heterocyclic ring size, and substantial progress is said to have been made. In view of the challenging scope of synthesis, and the importance of this project to the program, it will be important for the project leader to closely monitor and direct as needed, the work that is proposed to be conducted in Dr. Hammer's laboratory. This project can conceivably have far reaching and substantial impact on nucleic acid-based technologies beyond the scope of the program.

Intrinsic Scientific Merit Score: 139

Personnel:

Name: Donald E. Bergstrom
Degree/Discipline/Date: Ph.D., organic chemistry, 1970
Role/Percent Effort: Project Leader, ten percent
Qualifications/Experience: Professor of Medicine and Chemistry at Purdue University since 1989 and Deputy Director of the Purdue Cancer Center since 1992; he has served on the editorial board of Nucleosides and Nucleotides since 1992 and has a productive publication record in synthetic aspects of that field.

Assessment in designated role: highly qualified.

Name: Robert P. Hammer
Degree/Discipline/Date: Ph.D., organic chemistry, 1990

Role/Percent Effort: Co-investigator, ten percent during nine month academic years/66.7 percent during three month summer
Qualifications/Experience: Postdoctoral 1990-1992 at ETH with world-renown Professor Albert Eschenmoser on project-related chemistry, and Assistant Professor of Chemistry at LSU since August 1992; he has only one relevant publication in J. Org. Chem. in 1987 and several proceedings abstracts.

Assessment in designated role: qualified

Name: Guanygi Wang
Degree/Discipline/Date: Ph.D., organic chemistry, 1987
Role/Percent Effort: Postdoctoral Fellow, 100 percent
Qualifications/Experience: Postdoctoral positions at the University of Arizona, University of Maryland, and Purdue University, the latter since 1989 with Dr. Bergstrom; he is listed as a co-author on about a dozen organic chemistry journal publications since 1987.

Assessment in designated role: qualified.

Name: Peiming Zhang
Degree/Discipline/Date: Ph.D., organic chemistry, 1980
Role/Percent Effort: Research Scientist, 50 percent.
Qualifications/Experience: Postdoctoral since 1990 with Dr. Bergstrom; he is co-author of only one publication, which is with Dr. Bergstrom on the synthesis of a relevant nucleotide analogue.

Assessment in designated role: qualified.

Name: Melissa Cothern
Degree/Discipline/Date: B.S., chemistry, 1993
Role/Percent Effort: Research Assistant, 100 percent at no cost
Qualifications/Experience: No publications

Assessment in designated role: no proven experience.

Budget: The requested budget is modest and is approved.

Assessment: Level of merit; outstanding.

Project 4: Engineering an Improved Thermostable Ligase
(Francis Barany, Ph.D.)

Description: (Applicant's description) One of the fundamental problems in detecting cancers in tissue samples is the need to distinguish a few cells containing the cancer mutation from the vast majority of normal cells. We have developed a novel polymerase chain reaction/ligase detection reaction method (PCR/LDR) for high throughput, low sensitivity mutation detection (1 in 10 2 to 10 3), and a PCR/restriction endonuclease/LCDR (PCR/RE/LDR) method for high sensitivity mutation detection (1 in 10 6 to 1 in 10 7, see Project 1 and 2). The enzyme which provides the specificity for these methods is Tth ligase, the gene for which was originally cloned in our laboratory. The limit of detection of these two methods would be significantly improved by increasing the specificity of Tth ligase.

We are developing a comprehensive approach to understanding the mechanism of Tth ligase action, and improving its fidelity for discriminating perfectly matched from mismatched substrates. The three parts to this program are: (i) Developing a rapid assay to test different reaction conditions, mutant Tth ligases, and demonstrates higher sensitivity then the discriminating base is on the 3' end of the test primer. Introducing a nucleotide analogue (see Project 3) or mismatched base adjacent to or near the discriminating base may increase the specificity of this reaction. Such modified oligonucleotide primers will be tested in our fidelity assay using wild type and mutant Tth ligase. (ii) Determining the 3-dimensional structure of Tth ligase-DNA complex. This structure will help to reveal the mechanism of DNA ligation, and provide an understanding of the specificity of the enzyme for mismatches at the nicked site. (iii) Using site-specific mutagenesis to construct mutant Tth ligases. Design of these mutants will be based on protein sequence homology and protein-sugar-phosphate backbone contacts as determined from the X-ray structure. We have already isolated and partially characterized over 30 site-specific Tth ligase mutants, and these will be tested in our fidelity assay.

Critique: The overall goal of this project is to better understand the functional domains of Tth ligase and its mechanism of action which will hopefully result in an enzyme with improved specificity. Although DNA ligases from both prokaryotes and eukaryotes have been studied for many years, there has not been a comprehensive attempt using modern methods of site directed mutagenesis to study their reactions.

The primary reason for undertaking this program is the LCR and a variation LDR which hold great promise as diagnostic methods for detection of specific mutations. The project leader is one of the inventors of LCR. Thermostable ligases give superior performance in LCR and if the specificity of the ligase could be increased the sensitivity of detection of mutations could be increased. In as much as LCR or variations of it are central to the cancer detection studies proposed in this program, this section is central to the entire program since increasing the sensitivity of LCR and LDR is seminal, particularly if multiplexing is to be done.

The construct of new mutants are based on the previous mutagenesis studies by the project leader. These studies have led to identification of the adenylation site of Tth ligase in the vicinity of residue 118. These studies confirm a conserved motif, KVDG, in Tth ligase, suggested by the other workers, to be important in adenylation. Other residues are also identified in these studies which are necessary for ligase activity but not for adenylation or deadenylation. The ability of the cloned thermostable Tth ligase to complement an E. coli host containing a temperature sensitive mutant ligase permits a very nice in vivo assay for functionality. Also, the thermostability of the Tth ligase makes possible a relatively simple partial purification procedure for the ligase which permits in vitro characterization of the reactions of the mutant enzymes. Finally, a sensitive fluorescent assay for ligase fidelity has been developed using an Applied Biosystems (ABI) DNA sequencer employing the ABI Genescan software which permits the detection of products in the range of 100 attomoles. Preliminary studies have already led to the interesting observation that Tth ligase shows greater specificity for perfectly matched substrate over mismatched substrate when the mismatch is on the 3' side of the nick.

Rationale for the design of mutants, particularly in the absence of a structure, is a key issue. At this time, the project leader has only sequence comparisons and consideration of conserved amino acids to go on. It is noted that no consideration seems to have been given if large numbers of mutants were to be

screened. This might only be possible by genetic means. Considering the biochemical assays available, which make use of the temperature stability of Tth ligase, only up to several thousand of mutants could be screened in a reasonable time but not millions. However, the project leader's success on two recently isolated mutants that show improved specificity for TG mismatches suggest that it is not too difficult to isolate relevant mutants. In any event, this is not considered a serious criticism since the Site Visit Team has a great deal of confidence in the project leader's ability to recognize and deal with this problem if it occurs.

Overall, the proposed site-directed mutagenesis studies are excellent. Clean and sensitive assay methods have been developed that should, at the very least, lead to much important information about the functional domains of Tth ligase and appear to have a good probability of yielding mutants with increased specificity for at least some mismatches.

The second Specific Aim entails the testing of modified oligonucleotides for improved specificity during ligation. The approach is to test the effect of these modifications on the specificity of ligation when they occur adjacent to and one base over from the discriminating base. If modifications can be found that increase this specificity, then potentially the sensitivity of LCR/LDR, and hence their diagnostic value would increase. To be tested are base transposing agents called "convertides" which are modified so that they have ambiguous hydrogen bonding properties which permits them to base pair with more than one base and in some cases with each of the natural DNA bases (A, T, G and C). The project leader appears to be interested in 1-(2'-deoxy- -D-ribofuranosyl)-3-nitropyrrole, abbreviated Q2, which can pair with all of the natural bases. It is noted that many other "convertides" are proposed to be made in Project 3 although it is not clearly stated whether any of these other than Q2 will be tested in this project. These studies rely on supply of the analogues by Drs. Bergstrom and Hammer, from Project 3. At the Site Visit, it was reported that Dr. Bergstrom had just delivered to the project leader a new modified nucleotide, Q18, not discussed in the application, which will be tested.

The approach taken by the project leader in studying these particular types of modified nucleotides is derived from the observations of others that allelic specific PCR amplification is enhanced by destabilizing bases near the 3'end of a primer. Most notable here is perhaps the work of Dr. Thilly at MIT in developing "MAMA" technique (MAMA = mismatch amplification mutation assay). The project leader states that "This concept of destabilizing the enzyme-nucleic acid complex to significantly increase the specificity of a reaction is the basis behind our site-specific mutagenesis studies [discussed in several sections of the application]". In the reviewers's opinion, the project leader makes a large leap in coming to this conclusion. Dr. Thilly's work used one or two natural bases at the 3' end of a PCR primer which were mismatches of bases on the template. This is quite different from the "convertides" to be used here which have ambiguous hydrogen bonding properties that allow them to pair to more than one of the natural bases. There is no way to predict the effect of these "convertides."

The general approach of testing the effect of modified nucleotides on ligation specificity is strongly supported. However, the proposed approach might be too narrowly focused on the Q2 nucleotide and the specific observation mentioned above that destabilizing bases near the 3' end of a primer enhances allelic specific

PCR amplification: It is difficult to predict the effects of almost any modified nucleotide. This section of the project might be strengthened by a broader approach including types of nucleotide modifications other than the "convertides" proposed. At the very least, "convertides" other than Q2 should be incorporated into these studies as soon as possible. Q2 was designed to be a universal nucleotide and a part of this design was to make it lack a hydrogen donor site and maximize stacking interactions (see Project 3). All of the other "convertides" proposed have a different design principle in that they all have a hydrogen donor site. It is not possible to predict how important this difference might be. In the reviewer's opinion, it should be addressed as soon as possible.

Determination of the three dimensional structure of Tth ligase is clearly an important problem. The actual work will be done by Dr. Aneel Aggarwal who has the experience to perform this study. No one can predict the difficulty in determining a protein structure. Even if the structure is not solved, this project can yield important information. However, the value of the information would increase greatly if the structure were solved, especially early on.

Intrinsic Scientific Merit Score: 145

Administrative Note: Attention is drawn to the potential scientific overlap between the proposed studies by Dr. Francis Barany and those funded at Applied Biosystems Inc., on "Ligation Amplification Technology" (2/1/92-1/31/97).

Personnel:

Name: Francis Barany
Degree/Discipline/Date: Ph.D., microbiology, 1981
Role/Percent Effort: Project Leader, ten percent
Qualifications/Experience: Dr. Barany was a postdoctoral fellow with Dr. Hamilton Smith at The Johns Hopkins from 1982-1985. He joined the faculty of Cornell University Medical College in 1985 as an Assistant Professor and currently holds the rank of Associate Professor. He is the author of numerous publications relevant to this Project. His research accomplishments include cloning of Tth ligase; he is one of the inventors of the Ligase Chain Reaction.

Assessment in designated role: highly qualified

Name: Aneel K. Aggarwal
Degree/Discipline/Date: Ph.D., biophysics, 1984
Role/Percent Effort: Co-Investigator, ten percent
Qualifications/Experience: Dr. Aggarwal was a postdoctoral fellow with Stephen Harrison at Harvard university. In 1989, he became Assistant Professor of Biochemistry and Molecular Biophysics at Columbia University. Dr. Aggarwal's postdoctoral research on the structure of phage 434 repressor-DNA complex and his more recent work since coming to Columbia University on the structure of BamHI, appear to make him well suited to lead the structural studies outlined in this project.

Assessment in designated role: well qualified

Name: Jianying Luo

Degree/Discipline/Date: Ph.D., biochemistry, 1992
Role/Percent Effort: Research Associate, 50 percent
Qualifications/Experience: has constructed, isolated, sequenced, and characterized proteins from the mutants of Tth ligase gene. Furthermore, she has performed all the ligase fidelity assays described in the preliminary results.

Assessment in designated role: qualified

Budget: appropriate as requested.

Assessment: Level of merit; excellent.

Project 5: Design and Synthesis of DNA and PNA Arrays
(George Barany, Ph.D.)

Description: (Applicant's description) The goal of this program project is to develop methods for identifying multiple gene mutations in cancers. For maximum utility, these methods must be able to recognized and discriminate between dozens or hundreds of mutations.

To accomplish this, we propose to capture specific ligase detection reaction (LDR) products on a spatially addressable array, such that the position of a signal identifies a mutation. Each LDR product will have a "zip code" tail, which will be selectively captured by a "complementary zip code" on a solid support. The complementary components can be DNA oligonucleotides or peptide nucleotide analogues (PNA). PNA/DNA hybrids have significantly higher Tm values than DNA/DNA hybrids. Incorporation of the nucleotide analogue, 5-propynyluridine, into DNA zip code and PNA address sequences will further increase and optimize Tm values (Project 3). Unreacted LDR primer may therefore be washed away at high temperatures allowing for a higher sensitivity in detecting LDR products. A reusable, universal addressable array could be used for detecting a wide range of cancer mutations, genetic diseases and infectious agents.

Implementation of these concepts, with the ultimate goal of achieving reliable and efficient materials and procedures that can be incorporated into easy-to-use, automated, low-cost diagnostic devices, will follow these aims: (i) Development and evaluation of solid support materials compatible with chemical synthesis of Dna oligonucleotides and PNA oligomers, and compatible with subsequent hybridization reactions. Surfaces, beads, or membranes will be functionalized , and extended as needed with hydrophilic spacers such as heterobifunctional polyethylene glycol (PEG) and/or carbohydrates. Chemistry for linking oligomers to the solid support, and/or solid-phase assembly of oligomers, will be developed. (ii) Establishment of methodology for synthesis of spatially addressable arrays of DNA oligonucleotides and PNA oligomers. Appropriate masking technology will expose defined regions of the solid support for attachment of pre-formed oligomers, or for chain elongation to assemble the needed oligomers. In the latter mode, segment condensation will be used when possible in order to provide efficient convergent synthesis, and because chemical "failures" will become "invisible" during the subsequent hybridization. (iii) Demonstration of scope and limitations of zip code concepts. As aims (i) and (ii) come to fruition, testing will be carried out (Core B). Design of primer and zip code structures will be facilitated by the informatics collaboration (Core A).

Critique: In order to carry out large-scale screening of mutations, this project proposes to develop spatially addressable arrays of oligonucleotides or peptide nucleotide analogs. Ligase reaction products from the technology devised in Projects 1 and 2, will be constructed with fluorescent groups, and will bear specific "zip-code" tails. The tails will be selectively captured by complementary zip-code probes immobilized on the array. Each zip-code will map for a specific known genetic mutation.

The concept of zip-codes, analogous to the multiplex detection scheme of Church and Gilbert, has been very well developed in this study. The codes are conceived to be unique 24-mers, designed to have very little similarity overlap. A novel block synthetic scheme, suited to this design is proposed for PNAs. 36 of the 256 different tetramers have been chosen on the basis of minimal similarity, and will be synthesized as building blocks for 24-mer arrays. Approximately 2 X 10 9 24-mers could be constructed from this building block set, out of a possible approximately 3 X 10 14 24-mers. Initially, five tetramers have been chosen as synthetic targets for feasibility testing of the array technology.

Before actual arrays are to be constructed, a significant amount of DNA/PNA array development work is proposed. First, solid supports compatible with DNA or PNA synthesis will be screened. A variety of potentially compatible materials will be investigated, including glass, plastic, cellulose, PEG-PS beads and a variety of membranes. The Project Leader has a good command of solid-phase synthesis techniques, brings strength to this part of the project, and has put together an impressive list of surface functional and linker groups. In addition to the compatibility with synthesis, the hybridization compatibility of theses supports will also be screened.

The investigators propose to adapt newly developed PNA chemistry to the 24-mer array format. Although PNA/DNA complexes are known to be extremely heat stable, it is not clear that this will actually help the detection discrimination issues presented in this proposed study. 24-mer DNA complexes are already quite stable, and the array design itself preselects optimal 24-mers for minimal cross-hybridization. A significant amount of chemistry must accompany the PNA array development, and although synthetically efficient and scientifically interesting, the central need for PNA arrays for achieving the goals of the program project remains unclear.

The proposed study schematically illustrates a microchannel device to construct the DNA or PNA arrays. Embedded in the actual fabrication of the arrays, is an enormous amount of engineering and development work; however, no workplan or budget is provided. During the site visit, the investigators deferred this development activity to collaborators in laboratories at Sirrus or Millipore. Unfortunately, the reviewers have no way to judge the commitment or capability of the investigators at Sirrus or Millipore to meet the priorities of this program project.

In summary, the strengths of this project lie in the novel investigations related to of DNA zip-codes. The concept is well developed, designed and a synthetic procedure for efficient construction of PNA arrays is proposed. The solid-phase chemistry in this section is outstanding. However, the lack of expertise and experimental planning in hybridization, fluorescence detection, and perhaps most importantly, the inability to construct the arrays seriously weakens this section.

Intrinsic Scientific Merit Score: 244

Personnel:

Name: George Barany
Degree/Discipline/Date: Ph.D., biochemistry, mathematics, organic chemistry, 1977
Role/Percent Effort: Project Leader, ten percent
Qualifications/Experience: Dr. Barany is currently Professor of Chemistry at the University of Minnesota. He an expert on new methods of peptide synthesis, orthogonal protection, organosulfur chemistry and various polymer functionalization chemistries. He has an outstanding publication record, and is considered to be highly qualified to carry out the synthetic and solid-phase chemistry on this project.

Assessment in designated role: highly qualified.

Name: Josef Vagner
Degree/Discipline/Date: Ph.D., biochemistry, 1990
Role/Percent Effort: Postdoctoral Associate, 100 percent
Qualifications/Experience: research worker in Prague (1986-1992) with experience in solid-phase and solution peptide synthesis, immunology, and protein chemistry. Presently, a postdoctoral fellow at the University of Minnesota.

Assessment in designated role: qualified.

Budget: The modest budget, as requested, is recommended for approval.

Assessment: Level of merit; very good to good.

Core A: Informatic Support for Cancer Detection Methods
(Neil R. Hackett, Ph.D.)

Description: (Applicant's description) The goal of this program project is to develop techniques that detect multiple cancer mutations, ultimately for the purpose of researching the relationship between genetic alterations and tumor behavior, and applying these techniques in clinical situations. Managing a database of cancer-associated mutations, developing multiplex assays for them and correlating multiple cancer mutations with disease outcomes will require a sophisticated level of data management.

Core A will provide informatics support for cancer detection according to the following aims: (i) Create and maintain a database of mutations associated with cancer, patient history and experimental results. A relational Client/Server database will be created on a central facility consisting of a SPARC station 10 running the Sybase database management system. Both published reports and results from Projects 1 and 2 will be collected. (ii) Analyze database for correlations of point mutations with clinical outcome. The significance of the cancer detection experiments in Projects 1 and 2 will be assessed by performing multivariate analysis on given mutations to determine whether they predict clinical outcome. (iii) Write programs for the choice of primers for PCR/LDR, LDR/PCR and PCR/RE/LDR protocols. The programs will be written in C language for use on IBM/PC or Macintosh computers with a simple text-base interface. Thee programs will aid primer design and calculation of modified primer Tm values for Projects 1, 2 and 3. (iv). Assist

in the design and analysis of oligonucleotide arrays for mutation detection. For project 5, potential schemes for array design will be explored to ensure arrays of the maximum difference in sequence between every pair of zip codes while maintaining a constant melting temperature. (v). Assist in the programming of the instruments in the diagnostics and evaluation core and interface these with the central database (for Core B). The informatics core will maintain a Client/Server database on a SPARC station, and provide programming support which is accessible to all participants in the program project. This research may lead to correlations between molecular markers and prognosis for lung, colon, breast, and cervical cancers.

Critique: This critique is based on the assessment of the investigators rather than their written or verbal description of the proposed work, since the core is vague in a number of areas.

The investigators have experience in biological sequence analysis, and together with Dr. Barany's knowledge of primer chemistry, should be well equipped to develop primer selection programs (task 3). The reviewers at the site visit asked Dr. Hackett specific questions about his strategy for this problem and received a response indicating Dr. Hackett's excellent command of this problem.

The staff has experience with instrumentation and should be able to program and interface the various instruments required by the project (task 5). However, The staff in this Core have a very limited or lack of experience in data base which is of concern to the reviewers. The application does not provides details on any aspect of the database design, not even examples of specific data that will be stored in the database, e.g., specific items of clinical history that will be stored. At the site visit, Dr. Hacket presented an example purporting to illustrate his approach to clinical data. The example was unconvincing from a clinical standpoint in that it contained no relevant clinical data. The computing aspect also revealed a serious technical flaw. (The table on the upper right hand corner contained data values that were used as column headings in the table in the lower left, violating a standard tenet of relational database design).

There is no technical discussion of the decision to use SYBASE. The proposal mentions the importance of client/server architecture which is valuable in many situations, but it is not clear that it is important in this study. SYBASE is a complicated product, and the staff appears to have no experience with it. Perhaps a simpler MaC-based product, such as Fourth Dimension, would be a better choice.

It is unclear as to why this project needs to maintain a database regarding cancer causing mutations, except for the mutations being analyzed in Projects 1 and 2. The project leaders of Projects 1 and 2 are well aware of the literature and on-going work related to the genes being studied in these projects, and are unlikely to require a database of all cancer causing mutations. The investigators may want to do a literature search to see if they have missed any new reports of relevant mutations, but this is best done by a search of MedLine and public databases using standard software.

The complex issues involved in clinical data management were not discussed at the site visit. Dr. Osborne volunteered that he and his co-workers at Strang have extensive experience in this area and would be delighted to help out. Although

this is viewed positively, this involvement should have been explicitly described in the application.

No technical discussion on how the Strang clinical database will be translated to the proposed database. It is not clear if the proposed database will use the same schema as Strang. If not, some translation effort is needed to bridge the differences.

Additionally, how data would be moved from Colorado to the proposed database is also unclear. Is the Colorado data is available on computer already, then translation issues similar to those mentioned above must be addressed. If not, substantial effort will probably be required to extract and computerize this information based on patient records.

Experience in Biostatistics area is yet another concern. The investigators provide no details on biostatistical aspect of the clinical design, nor is there any discussion of statistical methods that will be used to analyze the data. At the site visit, Dr. George Wong, a bio-statistician at Strang, expressed his enthusiasm for the proposed work, but confirmed that he has not been involved in the project to date. The involvement of a biostatistician is essential for the design of the proposed clinical studies and the analysis of the results.

Project 1 includes some discussion of data management and statistical analysis for data collected in Colorado. There seems to be no coordination between the Colorado effort and the Informatics Core. These two efforts propose to use different types of computers (Mac versus Unix), different database management systems (Paradox versus SYBASE), and different statistical packages (SYSAT and SAS versus SPSS).

Personnel:

Name: Neil R. Hackett
Degree/Discipline/Date: Ph.D., biochemistry, 1982
Role/Percent Effort: Core Leader, 20 percent
Qualifications/Experience: postdoctoral training in the Department of Chemistry at MIT; Assistant Professor in the Department of Molecular Biology at Vanderbilt University for three years, and since 1989 has been Assistant Professor in the Department of Microbiology at Cornell University. Since 1990, also served as the Manager of the Molecular Biology Computing Facility at Cornell University. He has several publications in Molecular Biology, but lists no publications related to informatics.

Assessment in designated role: qualified.

Name: Aaron Giles
Degree/Discipline/Date: B.S., physics, 1992
Role/Percent Effort: Programmer, 100 percent
Qualifications/Experience: worked as a programmer at the High Energy Physics Department, University of Chicago through June of 1993 at which time, he joined the Office of Academic Computing at the Cornell University. His CV suggests experience in software development at the hardware/software boundary, such as firmware for attached laboratory devices, print drivers, and image viewing software.

Assessment in designated role: qualified.

Budget: The operating system ($5,000) is included in the "Equipment" category and is thus deleted from the "Supplies" category. The budget of $2,500 in the "Other Expenses" category is not justified and therefore deleted. Due to the excessive requested budget in the years 02-05, the costs for "Equipment" (-$2,000), and for maintenance in the "Supplies" categories (-$1,000) are recommended to be reduced. Additionally, the cost for "Other Expenses" (-$2,500), is recommended to be deleted as it is not justified.

Assessment: Level of merit; acceptable.

Core B: Instrumentation and Mutation Detection
(Francis Barany, Ph.D)

Description: (Applicant's description) Correlations of multiple cancer mutations with disease outcome will require the ability to perform high throughput mutation detection. The goal of this core is to provide the instrumentation and mutation detection support required to achieve large scale identification and analysis of mutations. Core B will work closely with Core A, the informatics support for cancer detection methods.

This Core will have the following responsibilities: (i) Providing instrumentation for oligonucleotide synthesis and analysis of cancer causing mutations. The PCR/LDR, LDR/PCR, and PCR/RE/LDR experiments described in Projects 1 and 2 require synthesizing large numbers of oligonucleotides. The products from these cancer detection amplifications will be separated and quantified on an ABI 373A DNA sequencer. By the third year we plan to automate some of the PCR/RE/LDR steps using a robotics workstation. (ii) Testing the efficiency and polymerase fidelity of nucleotide conversions using convertide oligonucleotides. The PCR/RE/LDR cancer detection scheme is dependent on the fidelity of thermostable polymerase extension off primers containing a 3' nucleotide analogue (Project 3). Using an assay we developed, the Core will test both the efficiency and fidelity of different polymerases for each base conversion. (iii) Testing oligonucleotide or PNA addressable arrays for quantitative cancer mutation detection. Large scale detection of a multitude of mutations will require addressable arrays. Mutations will be distinguished by the position of a fluorescent signal on the array. The Core will test arrays synthesized in Project 5 for fluorescent detection of LDR and LCR oligonucleotide products, using a Molecular Dynamics FluorImager 575.

Critique: Core B is considered as an essential part of the program project. The specific responsibilities include providing oligonucleotides to the rest of the program, robotics support for the PCR/RE/LDR assays, testing the efficiency and polymerase fidelity of nucleotide conversions, and testing of the oligonucleotide or PNA arrays for quantitative cancer mutation detection.

With respect to oligonucleotide synthesis, an upgraded Model 394 automated 4-column synthesizer donated by ABI is available for producing the required oligonucleotides. This is now conventional technology which should not pose any difficulties. The investigators in this group have experience with gel-purified oligonucleotides, which represents valuable know-how as purity of probes and primers will be a critical issue.

Currently, products from the experiments described in projects 1 and 2 are separated and quantified on a Model 373A DNA sequencer which has also been donated by ABI and is more or less turn-key technology. The concept of using, for example, hexaethylene oxide "tails" to differentiate multiplexed PCR/LDR and related amplification products has already been proven in referenced publications 2 and 3. There is ample precedent for the use of capillary electrophoresis to achieve the proposed separations. The proposed "zip code" concept for the PNA arrays should be an effective detection scheme if implemented into arrays. The Beckman Biomek or ABI Catalyst 800 robotics equipment for pipeting, etc. seems reliable and justifiable.

In this Core, the efficiency and polymerase fidelity of nucleotide conversions using convertide oligonucleotides provided by Project 3 will be tested. The applicant has devised a series of assays to determine how well a nucleotide analogue can "read" a natural base, which were considered sufficient by the review committee.

Finally, oligonucleotide or PNA addressable arrays for quantitative cancer detection will be tested. Test arrays are to be provided by Project 5 and hybridization will be detected using a Molecular Dynamics FluorImager 575 to be purchased in year-1 for $80,000. This system will likely require a substantial amount of unbudgeted software development. The hybridization assays will also require a substantial amount of optimization, especially if quantitation is desired. Overall, the core group could benefit from expertise in software and instrumentation development.

Personnel:

Name: Francis Barany
Qualifications described earlier under Principal Investigator and Project 4.

Assessment in designated role: highly qualified.

Name: Matthew Lubin
Qualifications described earlier in Project 2.

Assessment in designated role: well qualified.

Name: Jianying Luo
Qualifications described earlier in Project 4.

Assessment in designated role: qualified.

Budget: The modest budget is approved as requested, except that the level of effort for the research technician is recommended to be changed from six months at 100 percent effort to one year at 50 percent effort.

Assessment: Level of merit; excellent.

Core C: Administrative Core
(Francis Barany, Ph.D.)

Description: (Applicant's description) The structure of the administrative core has been designed to meet the planning and managerial needs of the program project. It is also aimed at combining the expertise of all institutional administrators and scientific program directors in order to enhance the proficiency of the research level in order to promote cooperative efforts at all collaborating institutions.

The administrative responsibility of the Administrative core will include: 1) monitoring the scientific and administrative diligence of each of the component projects and re-allocating resources when and if necessary, 3) scheduling meetings of group investigators to be held approximately three time a year, 4) to keep the NIH Scientific Program Director/Coordinator apprised of group progress, changes in scientific aims, personnel, etc. 5) preparing progress reports for the NIH 6) assuring that core resources are providing adequate support to all projects, 7) coordinating group activities with all external companies in keeping with the formal collaborative research arrangements, and 8) monitoring inventions and invention disclosures supported by the program. In this fashion, the Administrative core will enhance maximum coordination between investigators participating in the program and facilitate the research objectives of the program project.

Critique: Dr. Barany's past research accomplishments demonstrate his ability to administer effectively individual R01 projects. Although his record of administrative experience other than that required for individual projects is only recent and limited, his success to date in assembling the program project key investigators is evident. His collaboration since 1991 with Dr. Wilson has resulted in a co-authored publication. He successfully initiated research discussions with Drs. Aggarwal, Hackett, and Lubin that led to the plans proposed as Project 2, Project 4, Core A, and Core B. He recruited Dr. Wilson for involvement in this program project. He has secured the commitment of nine potential members of an external advisory panel.

The plan to conduct productivity evaluations at formal annual meetings is appropriate. Indicators to monitor progress and expenditures are not established. Decision to date appear to rely on his personal communications with others. Effective interchange requires a structure commitment to frequent convening of the collaborators as a group to discuss current considerations; however, the absence of a plan for such sessions between annual meetings raises a serious concern.

A process for allocation and reallocation of funding cited as a responsibility within this Core is not documented. A role for the Administrative Core in contributing to the management of resources at the project level is not defined. The program appears to rely largely on effective administration at the level of project and core leaders, with the Administrative Core responsible primarily for interproject communications and reporting.

The issue of potential changes in leadership presents some concerns. A single successor to Dr. F. Barany as the Principal Investigator is not obvious. In the absence of published work in this field, it is not apparent that Dr. Lubin (Project 2) nor Dr. Aggarwal (Project 4) could conduct these projects as proposed in his absence. The plan to include external advisors in internal advisory meetings is

commendable. The application provides letters of commitment from the proposed external advisors, although these letters do not explicitly confirm the expectation to attend an annual meeting. To maximize effectiveness of the annual review, plans are needed regarding preparations for this session and follow-up on subsequent recommendations. The numerous letters of support from other individuals interested in development of the program actually raise concerns about the focus and priorities of the planned studies.

Personnel:

Name: Francis Barany
Degree/Discipline/Date: Ph.D., microbiology, 1981
Role/Percent Effort: Core Leader, five percent
Qualifications/Experience: Dr. Barany is personally involved in the current research collaborations between Project 2 and 4, and Cores A and B. He personally has secured the commitment of potential members of the external advisory panel. Dr. Barany's biographical sketch does not indicate any administrative experience to date beyond leadership of an R01 grant. His recruitment of an experienced administrator as Co-Leader of the Administrative Core shows good judgment.

Assessment in designated role: qualified

Name: Michael J. Bunk
Degree/Discipline/Date: Ph.D., nutritional biochemistry, 1980
Role/Percent Effort: Core Co-leader, 5(YR-01) to 10(YR-02-05)
Qualifications/Experience: Director, Research Resources Management, Strang Cornell Cancer Prevention Center since 1993; previously Assistant Program Director, Clinical Nutrition Research Unit (1987-89), Director of Foundation Relations (1989-91), and Senior Grants Management Specialist (1991-93) at Memorial Sloan-Kettering. Dr. Bunk has served for one year, at an affiliate (Strang) of the applicant organization, in a title very similar to the proposed role. The application, however, does not describe his current duties nor the relevance of this experience to his proposed role. He has a total of six years of potentially relevant prior administrative experience in three different programs outside the applicant organization. The current application and the collaborations developed to date demonstrate administrative skill, but Dr. Bunk's specific contribution to date to these preparations to evaluate the percent effort requested.

Assessment in designated role: qualified.

Budget: A number of administrative systems will need to be developed and established immediately in Year -01. In the absence of evidence that the administrative responsibilities will increase dramatically from the Year 02, it is recommended that the Program Coordinator's effort should remain at 30 percent during this time. The remaining budget is appropriate as requested.

Assessment: Level of merit; good to acceptable.

Women and Minorities in Study Population: All of the proposed clinical materials will come from existing tissue banks from which individual patient identification cannot be made. Therefore this proposal is considered exempt from Women and Minority considerations.

BUDGET

Project 1 (Dr. Francis Barany, Cornell University Medical College

BUDGET CATEGORIES	Requested	Recommended
THIRD PARTY COSTS		
Direct	71,545	71,545
Indirect	28,959	28,959
TOTAL	100,504	100,504

Project 1 (Dr. FrancisBarany, Cornell University Medical College (The Children's Hospital))

BUDGET CATEGORIES	Requested	Recommended
PERSONNEL	54,450	54,450
EQUIPMENT	2,595	2,595
SUPPLIES	10,300	10,300
TRAVEL-Domestic	1,200	1,200
THIRD PARTY COSTS		
Indirect Costs		
42 percent	28,959	28,959
OTHER EXPENSES	3,000	3,000
TOTAL	100,504	100,504

Project 2

BUDGET CATEGORIES	Requested	Recommended
PERSONNEL	84,860	84,860
SUPPLIES	14,500	14,500
TRAVEL-Domestic	1,200	1,200
OTHER EXPENSES	3,000	3,000
TOTAL	103,560	103,560

Project 3 (Dr. Francis Barany, Cornell University Medical College

BUDGET CATEGORIES	Requested	Recommended
THIRD PARTY COSTS		
Direct	131,170	131,170
Indirect	60,888	60,888
TOTAL	192,058	192,058

Project 3 (Dr. Donald Bergstrom, Perdue University)

BUDGET CATEGORIES	Requested	Recommended
PERSONNEL	44,305	44,305
SUPPLIES	13,000	13,000
TRAVEL-Domestic	1,200	1,200
THIRD PARTY COSTS		
Indirect Costs 51 percent	33,663	33,663
OTHER EXPENSES	7,500	7,500
TOTAL	99,668	99,668

Project 3 (Dr. Robert P. Hammer, Louisiana State University)

BUDGET CATEGORIES	Requested	Recommended
PERSONNEL	40,800	40,800
EQUIPMENT	4,665	4,665
SUPPLIES	14,000	14,000
TRAVEL-Domestic	1,200	1,200
THIRD PARTY COSTS		
Indirect Costs 45 percent	27,225	27,225
OTHER EXPENSES	4,500	4,500
TOTAL	92,390	92,390

Project 4 (Dr. Francis Barany, Cornell University Medical College)

BUDGET CATEGORIES	Requested	Recommended
PERSONNEL	37,847	37,847
SUPPLIES	10,500	10,500
TRAVEL-Domestic	1,000	1,000
THIRD PARTY COSTS		
Direct	60,918	60,918
Indirect	40,815	40,815
OTHER EXPENSES	3,000	3,000
TOTAL	154,080	154,080

Project 4 (Dr. Aneel Aggarwal, College of Physicians and Surgeons of Columbia University)

BUDGET CATEGORIES	Requested	Recommended
PERSONNEL	45,918	45,918
SUPPLIES	10,500	10,500
TRAVEL-Domestic	1,500	1,500
THIRD PARTY COSTS	40,815	40,815
OTHER EXPENSES	3,000	3,000
TOTAL	101,733	101,733

Project 5

BUDGET CATEGORIES	Requested	Recommended
PERSONNEL	64,069	64,069
SUPPLIES	15,500	15,500
TRAVEL-Domestic	1,200	1,200
THIRD PARTY COSTS		
Indirect Costs	85,769	85,769
40 percent Direct	34,308	34,308
OTHER EXPENSES	5,000	5,000
TOTAL	120,077	120,077

Core A

BUDGET CATEGORIES	Requested	Recommended
PERSONNEL	70,488	70,488
EQUIPMENT	15,000	15,000
SUPPLIES	16,500	11,500
TRAVEL-Domestic	1,200	1,200
OTHER EXPENSES	2,500	0
TOTAL	105,688	98,188

Core B

BUDGET CATEGORIES	Requested	Recommended
PERSONNEL	60,638	60,638
EQUIPMENT	80,100	80,100
SUPPLIES	11,000	11,000
OTHER EXPENSES	5,000	5,000
TOTAL	156,738	156,738

Core C

BUDGET CATEGORIES	Requested	Recommended
PERSONNEL	22,836	22,836
TRAVEL-Domestic	9,000	9,000
OTHER EXPENSES	5,500	5,500
TOTAL	38,336	38,336

SUMMARY RECOMMENDED BUDGET*

BUDGET CATEGORIES	01 Year	02 Year	03 Year	04 Year	05 Year
Personnel	276,669	343,803	343,803	343,803	343,803
Consultant Costs	1,000	1,000	1,000	1,000	1,000
Equipment	95,100	6,000	76,000	6,000	6,000
Supplies	47,500	43,500	43,500	43,500	43,500
Travel (Domestic)	12,400	12,400	12,400	12,400	12,400
Third Party Costs					
Direct Costs	349,402	350,142	350,142	350,142	350,142
Indirect Costs	164,970	164,970	164,970	164,970	164,970
Other Expenses	16,500	16,500	16,500	16,500	16,500
TOTAL	963,541	938,315	1,008,315	938,315	938,315

*Appropriate escalation factors to be added at the time of an award.

NATIONAL CANCER INSTITUTE
SITE VISIT TEAM ROSTER
1 P01 CA65930-01, Francis Barany, Ph.D.
Title: New Methods For Cancer Detection
Cornell University Medical College
New York, New York
May 31-June 2, 1994

Revised 5/17/94

Chairman

John C. Ruckdeschel, M.D.
Professor of Medicine
Director, H. Lee Moffitt Cancer Center
Tampa, Florida

Reviewers

John W. Chase, Ph.D.
Scientific Director
United States Biochemical Corporation
Cleveland, Ohio

Steven P. A. Foder, Ph.D.
Scientific Director
Chief Technical Officer
Affymetrix, Inc.
Santa Clara, California

Suzanne Fuqua, Ph.D.
Assistant Professor
University of Texas Health Science Center
Department of Medicine/Oncology
San Antonio, Texas

Robert G. Gerlach, MPA
Administrative Cancer Center
Cleveland Clinic Foundation
Cleveland, Ohio

Nathan Goodman, Ph.D.
Associate Director, Center for Genome Research and Senior Research Scientist
Whitehead Institute
Cambridge, Massachusetts

Michael McClelland, Ph.D.
Director, California Institute o Biological Research
Lajolla, California

Christian Overton, Ph.D.
Research Associate Professor
Department of Genetics
University of Pennsylvania Schoo of Medicine
Philadelphia, Pa.

Jim G. Wetmur, Ph.D.
Professor
Department of Microbiology
Mt. Sinai School of Medicine
New York, New York

Gerald Zon, Ph.D.
Vice President, Medicinal Chemistry
Lynx Therapeutics, Inc.
Hayward, California

Scientific Review Administrator

Devi Vembu, Ph.D.
Scientific Review Administrator
Grants Review Branch
Division of Extramural Activities
National Cancer Institute
Rockville, Maryland

Program Representative

Sheila Taube, Ph.D.
Chief, Diagnosis Branch
Division of Cancer Biology, Diagnosis and Centers (DCBDC)
National Cancer Institute
Rockville, Maryland

Appendix 3

(To Declaration of Francis Barany under 37 CFR § 1.608(b))

CORNELL UNIVERSITY MEDICAL COLLEGE
Department of Microbiology, Box 62
1300 York Avenue, New York, NY 10021
Telephone: (212) 746-6509 Fax: (212)746-8587

November 11, 1994

Dr. James Jacobson
National Cancer Institute
Executive Plaza North
Rm 513
61-30 Executive Blvd.
Rockville, MD 20892
Tel.# 301-496-1591

Dear Jim,

I appreciate our frank discussion on Friday, October 28th, and am following up on your suggestion to send you a brief letter outlining the issues and concerns of the co-investigators of our program project grant entitled "New Methods for Cancer Detection." We are strong supporters of peer review, and have found review of our own proposal (not funded) both fair and very constructive. We are writing this letter because we would like to preserve the sanctity of NIH and NCI review.

An event of the last week has reopened our concerns that there were hidden conflicts with the review of our program project grant. One of our colleagues at the recent American Society for Human Genetics meeting reported that a group from Affymetrix, a for profit corporation of Santa Clara, CA, presented work wherein they are using DNA ligase to join oligonucleotides hybridized on a DNA chip. In particular they claim that perfectly matched hybridization units are covalently joined and those in which there is as little as a single base mismatch at the ligation junction are not. They claim this offers greater discrimination between perfect hybridization units and single nucleotide mismatches, because the unligated labeled probe molecules can subsequently be washed away.

It is our opinion the above presentation constitutes a substantial overlap with the contents and ideas presented in our NCI program project grant. This material was made available to the NCI site visit team, which included Dr. Stephen Fodor, Scientific director and Chief Technical Officer of Affymetrix Inc. on April 29th, 1994 in a 448 page document, as well as discussed in detail during a 6 hr site visit held at Cornell University on June 1st, 1994.

We were informed of the members of the site visit team via Fax only on April 29th, 1994, at which time grants were already sent to the reviewers. At the time, we expressed our concerns to Drs. Devi Vembu, David Irwin, Shiela Taube, and yourself, that certain members of the committee, among them Dr. Stephen Fodor, may have conflicts of interest. One such member was removed, but Dr. Fodor was not. You assured us that the NCI would vigilantly guard against any conflicts of interest problems.

During the eight months that we organized and wrote our NCI program project grant, we realized that DNA arrays (Project 5) would be an important part of the proposal. At the time of our grant submission in February 1994, we were aware that competing groups from Affymetrix and Beckman were developing DNA arrays, however no detailed description of their work appeared in the peer reviewed literature. Since that time, two reports on DNA arrays have appeared, Pease et al. (with S. Fodor as senior author), Proc. Natl. Acad. Sci. 91: 5022-5026 in May 1994, and Eggers et al. (HARC-Beckman group)

Biotechniques, 17:516-524 in September 1994. It is clear that the DNA arrays as described in the Affymetrix PNAS report would not be sensitive enough to detect cancer mutations as properly described in our program project grant.

Neither of these reports mentioned use of a DNA ligase for discriminating single base differences. Indeed, during the site visit one of us (F.B.) specifically addressed a question from Dr. Fodor, explaining to him that the specificity of a ligase enzyme will always give greater discrimination of a single base difference than hybridization alone.

The above concepts are key to reducing array based DNA diagnostics to practical use. As a direct consequence of one of our (F.B.) publications introducing LCR and LDR in the peer reviewed literature in 1991, ligase based detection methods have now become an area of intense competition between academic and industrial research teams. Affymetrix clearly has a substantial financial interest in this area, as evidenced by a $30,000,000 grant awarded by the National Institutes of Standards and Technology (NIST) in October 1994. The application for this award was likely prepared during the time of our program project site visit, since the NIST program was announced on April 25, 1994, and applications were due about two months later. (Affymetrix had access to our grant application from April 29, 1994 to June 1, 1994.) Our NCI proposal contained a substantial amount of confidential information relating to arrays and ligase detection technology. Based on their recent ASHG presentation, we believe the Affymetrix NIST proposal may have substantial overlap with the aims of our NCI program project grant proposal.

In order to determine if scientific ethics were breached, we feel it is appropriate that there be a thorough evaluation of the Affymetrix NIST proposal for overlap with our NCI program project grant. To avoid similar potential problems of scientific ethics when we resubmit our own NCI proposal, we feel it would be inappropriate for us to perform this thorough evaluation. We wish to request that either you or an impartial third party chosen by you conduct a comparative evaluation of the Affymetrix NIST proposal and our NCI program project proposal. In particular one should investigate whether Affymetrix proposed using thermostable ligase or borrowed intellectual and design concepts from our zip-code capture array. If there is overlap that involves plagiarism of intellectual property, then we would like a formal investigation.

We know that you and NCI are fully committed to maintaining the highest level of integrity during the review process and shall assist you in any way possible.

Sincerely yours,

Francis Barany
Associate Professor of Microbiology
Tel.# 212-746-6509

Donald E. Bergstrom
Professor of Medicinal Chemistry
Tel.# 317-494-6275

c.c. Dr. Shiela Taube
c.c. Dr. Robert Strausberg